



Imagerie et spectro-imagerie X appliquees a l'etude des proprietes du milieu intergalactique dans les amas de galaxies en cours de coalescence.

Herve Bourdin

► To cite this version:

Herve Bourdin. Imagerie et spectro-imagerie X appliquees a l'etude des proprietes du milieu intergalactique dans les amas de galaxies en cours de coalescence.. Astrophysique [astro-ph]. Université Pierre et Marie Curie - Paris VI, 2004. Français. NNT : . tel-00011275

HAL Id: tel-00011275

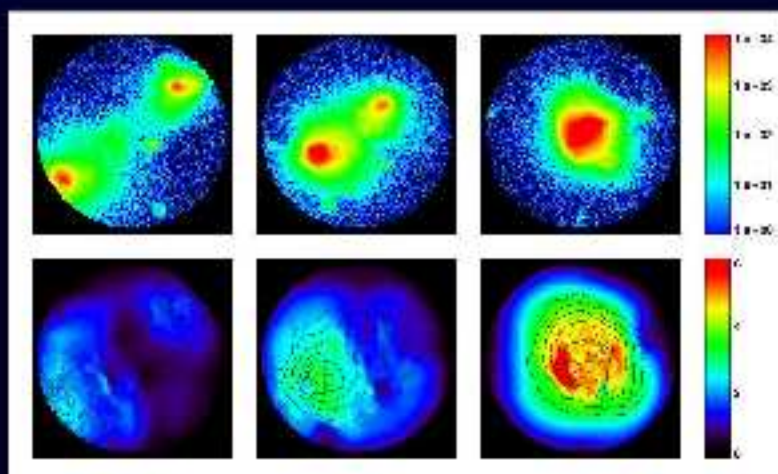
<https://theses.hal.science/tel-00011275>

Submitted on 2 Jan 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Imagerie et spectro-imagerie X appliquées à l'étude des propriétés du milieu intergalactique dans les amas de galaxies en cours de coalescence.



Hervé Bourdin

Thèse de doctorat présentée pour obtenir le titre de docteur en Astrophysique de
l'Université Paris 6, soutenue le 30 Mars 2004 à l'Observatoire de Nice devant le
jury composé de:

Patrick Boissé	Président du jury
William Forman	Rapporteur
Laure Blanc-Féraud	Rapporteur
Eric Slézak	Directeur de thèse
Albert Bijaoui	Co-directeur de thèse
Jean-Luc Sauvageot	Examineur

UNIVERSITE PIERRE ET MARIE CURIE
Ecole Doctorale d'Astronomie et d'Astrophysique d'Ile de France

Imagerie et spectro-imagerie X appliquées à l'étude des propriétés du milieu intergalactique dans les amas de galaxies en cours de coalescence.

Hervé Bourdin

Thèse de doctorat présentée pour obtenir le titre de docteur en Astrophysique de
l'Université Paris 6, soutenue le 30 Mars 2004 à l'Observatoire de Nice devant le
jury composé de:

Patrick Boissé	Président du jury
William Forman	Rapporteur
Laure Blanc-Féraud	Rapporteur
Eric Slézak	Directeur de thèse
Albert Bijaoui	Co-directeur de thèse
Jean-Luc Sauvageot	Examineur

Remerciements

Un travail de recherche est souvent un travail d'équipe. Celui-ci est avant tout le fruit de l'enseignement méthodologique et académique que j'ai reçu de mes directeurs de thèse, mais aussi de l'expérience que j'ai pu acquérir auprès de nombreux collaborateurs. Je tiens tout d'abord à remercier Eric Slézak et Albert Bijaoui de m'avoir proposé ce sujet de thèse et de m'avoir encadré durant ces trois années. Merci aussi à Jean-Luc Sauvageot avec qui j'ai beaucoup collaboré ainsi qu'à Monique Arnaud, Elena Belsole, Gabriel Pratt et Romain Teyssier, alors membres du Service d'Astrophysique du CEA à Saclay. Merci encore à Trevor Ponman, de l'Université de Birmingham, ainsi qu'à Sophie Maurogordato, Christophe Benoist et Chiara Ferrari de l'équipe "Analyse de données et Cosmologie" à l'observatoire de la Côte d'Azur.

Ce travail ayant abouti à la présentation d'une thèse, je remercie Laure Blanc-Féraud et William Forman d'avoir accepté de rapporter sur le manuscrit et d'avoir sensiblement contribué à l'améliorer. Merci aussi à Patrick Boissé d'avoir accepté d'en présider le jury.

De nombreuses personnes ont en outre indirectement contribué à l'aboutissement de ce travail. Merci à James Rich et Bertrand Chauvineau pour leurs cours de cosmologie et de relativité générale auxquels j'ai pu assister, ainsi qu'à Patrick Flandrin et Andréa Ferrari pour leurs cours de traitement du signal. Merci à Pierre Encrenaz et aux secrétariats de l'école doctorale d'Astronomie et d'Astrophysique d'Ile de France pour leur aide dans diverses démarches. Merci enfin à l'ensemble du personnel de l'observatoire de Nice avec qui j'ai pu travailler dans d'excellentes conditions et en particulier aux membres du service informatique ainsi qu'au personnel du restaurant.

Imagerie et spectro-imagerie X appliquées à l'étude des propriétés du milieu intergalactique dans les amas de galaxies en cours de coalescence.

Les amas de galaxies sont des surdensités de matière gravitationnellement liées remplies d'un gaz chaud et ionisé émettant un rayonnement X. Ils se forment durant des phases d'accrétion de sous-groupes, au cours desquelles le gaz subit des processus de choc et de mélange qui perturbent ses propriétés physiques à l'équilibre hydrostatique. Pour cartographier les distributions spatiales d'émissivité, de température et d'entropie du plasma intergalactique observé par les télescopes X, nous avons comparé différents algorithmes d'imagerie multiéchelle, puis développé et testé un nouvel algorithme de spectro-imagerie. Dans cet algorithme, le paramètre recherché est estimé à partir d'une statistique de comptage dans différents éléments de résolution spatiaux, puis ses variations spatio-fréquentielles sont codées par des coefficients en ondelettes de Haar. La distribution spatiale optimale du paramètre est finalement restaurée en seuillant la transformée en ondelettes bruitée.

Mots-clé

Cosmologie observationnelle – Formation des structures – Amas de galaxies – Milieu intra-amas
Astronomie X – Imagerie – Spectro-imagerie – Ondelettes

X-ray imaging and spectro-imaging techniques for investigating the intergalactic medium properties within merging clusters of galaxies.

Clusters of galaxies are gravitationally bound matter over-densities which are filled with a hot and ionized gas emitting in X-rays. They form during merging phases of subgroups, so that the gas undergoes shock and mixing processes which perturb its physical properties at hydrostatic equilibrium. In order to map the spatial distributions of the gas emissivity, temperature and entropy as observed by X-ray telescopes, we compared different multi-scale imaging algorithms, and also developed and tested a new multi-scale spectro-imaging algorithm. With this algorithm, the searched parameter is first estimated from a count statistics within different spatial resolution elements, and its space-frequency variations are then coded by Haar wavelet coefficients. The optimal spatial distribution of the parameter is finally restored by thresholding the noisy wavelet transform.

Keywords

Observational cosmology – Structure formation – Clusters of galaxies – Intra-cluster medium
X-ray astronomy – Imaging – Spectro-imaging – Wavelets

Table des matières

I	Formation et évolution des amas de galaxies.	15
1	Les amas de galaxies et l’astronomie X.	17
1.1	Les amas et les groupes de galaxies.	18
1.2	Le plasma intra-amas.	20
2	Les amas relaxés	23
2.1	Profils de densité.	24
2.1.1	Densité de matière baryonique.	24
2.1.2	Densité de matière noire	25
2.2	Calculs de masses.	25
2.2.1	Masse de gaz.	25
2.2.2	Masse de matière noire.	26
2.3	Relations autosimilaires.	27
2.3.1	Emissivité d’une sphère de densité moyenne constante.	27
3	Formation des structures de l’Univers.	29
3.1	Evolution dynamique de l’Univers	30
3.2	Formation des structures	31
4	Evolution des amas de galaxies dans le cadre de la formation des structures.	35
4.1	Processus de chocs adiabatiques	36
4.2	Simulations numériques d’amas en coalescence.	38
4.2.1	Simulation de coalescences bimodales.	38
4.2.2	Formation d’un amas massif au sein d’une simulation de formation des structures à grande échelle	39
II	Imagerie et spectro-imagerie X.	43
1	Le télescope XMM-Newton et les spectro-imageurs EPIC	45
1.1	Présentation du télescope	46
1.2	Nature du signal observé	47
2	La transformée en ondelettes.	51
2.1	L’analyse temps-fréquence et la transformée en ondelettes.	52
2.2	Analyse multirésolution	52
2.2.1	Sous-espaces d’approximation multirésolution	53
2.2.2	Fonction d’échelle	54
2.2.3	Relations inter-échelles et filtrage	55
2.2.4	Analyse multirésolution et transformée en ondelettes à deux dimensions	55
2.3	Algorithmes de calcul des transformées en ondelettes discrètes	57
2.3.1	Algorithme multirésolution	57
2.3.2	Algorithme à trous.	62

2.3.3	Opérateurs de filtrage	64
2.4	Applications: Compression de l'information et débruitage.	66
2.5	Choix d'une base d'ondelettes orthogonales.	66
3	Restauration d'image à faible statistique appliquée à l'émission X du plasma intra-amas.	69
3.1	Le signal: Emissivité du plasma intra-amas.	70
3.1.1	Du processus de Bernoulli au processus de Poisson	70
3.1.2	Une source étendue avec des variations spatiales.	70
3.2	Choix de l'algorithme	71
3.2.1	L'approche multiéchelle	71
3.2.2	Débruitage	74
3.2.3	Comparaison des algorithmes.	80
3.3	Restauration d'image adaptée aux processus de Poisson à faible statistique	80
3.3.1	Seuillage de la transformée.	80
3.3.2	Algorithme de restauration.	82
3.3.3	Construction de l'algorithme	84
4	Spectro-imagerie X appliquée à l'étude des propriétés du plasma intra-amas.	89
4.1	Les observables: les paramètres physiques du milieu intra-amas.	90
4.2	L'estimateur local $\hat{\theta}$ et la spectroscopie X.	90
4.2.1	Estimateurs de variance minimum limite (MVB) et borne de Rao-Cramer. .	91
4.2.2	Fonction de vraisemblance.	94
4.2.3	Méthode du maximum de vraisemblance.	95
4.2.4	Estimateur du maximum de vraisemblance et méthode des moindres carrés. .	95
4.2.5	Le rapport de dureté.	96
4.2.6	Application à la spectroscopie X.	98
4.3	Algorithme multiéchelle	105
4.3.1	Choix de l'algorithme.	106
4.3.2	Principe de l'algorithme.	106
4.3.3	Cartes d'erreur.	109
4.3.4	Implémentation de l'algorithme de spectroscopie.	109
4.3.5	Simulations d'observations.	112
4.4	Publication à propos de l'algorithme de spectro-imagerie.	114
4.5	Conclusion.	130
III	Applications	131
1	Restauration d'images	133
1.1	HCG 16 / Emission étendue dans un groupe compact.	134
2	Spectro-imagerie	137
2.1	Deux amas de galaxies bi-modaux en coalescence: Abell 3921 et Abell 1750	138
2.1.1	A3921	138
2.1.2	A1750	140
2.2	Ecart à l'isothermalité dans un amas massif: Abell 2163	141
IV	Annexes	149
A	Publications.	151
B	Notations	183

Introduction

Observé dans le domaine des longueurs d'onde visibles, le contenu matériel de l'Univers observable se présente sous la forme d'une distribution de galaxies homogène à grande échelle mais structurée en filaments séparés par de grands vides à plus petite échelle. Les galaxies elles-mêmes ne sont pas réparties de manière homogène dans les filaments mais s'agrègent en groupes de quelques unités. A l'intersection des filaments, on observe des amas de plusieurs centaines de galaxies. Cette composante visible de la matière n'étant pas assez massive pour assurer la cohésion gravitationnelle des structures observées, on suppose qu'elle ne fait que tracer la distribution d'une matière noire non collisionnelle dont la contribution en masse constitue l'essentiel de la quantité de matière contenue dans l'Univers. Les amas de galaxies sont des structures gravitationnellement liées dont la masse se répartit entre 80 % de matière noire et 20 % de matière baryonique. La matière baryonique contenue dans les amas comprend les étoiles et le milieu interstellaire constitutif des galaxies, mais aussi le plasma intra-amas, qui constitue un gaz chaud (10^8K), peu dense (10^{-3}cm^{-3}) et ionisé, remplissant l'espace intergalactique. Il émet à la fois un continuum de rayonnement de freinage des électrons et de raies d'ionisation d'éléments lourds dans le domaine des X.

Selon le modèle cosmologique admis actuellement, les structures de l'Univers, constituées essentiellement de matière noire et visibles sous la formes de galaxies, groupes et amas de galaxies, se forment hiérarchiquement sous l'action de la gravitation, par accréation de matière le long des surdensités filamentaires. Après la phase, dite linéaire, où les sous-groupes de matière noire non collisionnelle se forment hiérarchiquement, le processus de formation des structures se complexifie au cours d'une seconde phase, dite non-linéaire. La matière baryonique est accrétée dans le potentiel gravitationnel formé par les structures de matière noire. Sa nature collisionnelle fait entrer en jeu des processus non gravitationnels dans la dynamique de la formation des structures. Ces processus sont aujourd'hui observables dans les groupes et les amas de galaxies qui sont les seules structures encore en formation dans le cadre du scénario hiérarchique. C'est le plasma intra-amas qui en est principalement le siège. En effet sa masse contribue significativement à la masse de matière baryonique et sa nature est plus collisionnelle que celle des galaxies. Ce plasma, confiné dans le puits de potentiel de matière noire des amas, est en première approximation à l'équilibre hydrostatique, mais il subit des coalescences de sous-groupes de masse, de taille et de température distinctes qui perturbent fortement cet état d'équilibre. Ces coalescences entraînent des processus de mélange et le développement d'ondes de chocs qui augmentent la température et l'énergie interne du gaz. Par ailleurs le milieu intra-amas interagit avec les galaxies: les vents galactiques contribuent à modifier l'énergie interne et la composition chimique du gaz, tandis que l'évolution stellaire et la morphologie des galaxies est perturbée par les mouvements du gaz.

Les groupes et les amas de galaxies sont donc des laboratoires idéaux pour étudier, à travers la dynamique du plasma intra-amas, les processus non gravitationnels entrant en jeu dans le cadre de la formation des structures. L'étude de ces processus est aussi motivée pour quantifier les perturbations d'énergie interne, d'entropie et de température du gaz, qu'ils entraînent par rapport à un scénario considérant les amas comme des systèmes isolés et relaxés, à l'équilibre hydrostatique. C'est en effet sous cette hypothèse que la température du plasma intra-amas permet d'estimer la masse des amas, tandis que son émissivité X permet d'estimer la masse de gaz et la fraction de baryons. Sous l'hypothèse d'équilibre hydrostatique et en supposant que la fraction de baryons se conserve, la morphologie des amas de galaxies se comporte de manière auto-similaire pour des amas de différentes masses. Cette propriété remarquable implique une corrélation entre la température et la luminosité moyenne des amas, $L_X \propto T^2$. La formation des amas étant assez récente, l'évolution de la distribution de masse des amas en fonction de leur distance à l'observateur –donc du temps–, fournit des contraintes sur les paramètres cosmologiques et en particulier la densité de matière. Des mesures de masses d'amas à différentes distances sont ainsi effectuées sur de grands échantillons en utilisant la spectroscopie X. Il est important que ces mesures ne soient pas biaisées par les effets non gravitationnels entrant en jeu dans la dynamique du gaz. Un moyen de quantifier ces effets, le cas échéant variables en fonction du temps et du taux de coalescence des amas, est d'étudier les écarts aux lois autosimilaires luminosité-température et masse-température qu'ils impliquent.

Les variations spatiales d'émissivité, de température et de métallicité du plasma intra-amas peuvent maintenant être étudiées à des résolutions de l'ordre de la dizaine de seconde d'arc grâce aux spectro-imageurs équipant les satellites X de dernière génération (Chandra; XMM-Newton). Dès les premières observations d'amas en coalescence, ces satellites ont permis de mettre en évidence avec une acuité sans précédent des variations de température du plasma associées à des ondes de choc ou de compression (Markevitch, 2001), ou encore à des discontinuités de contact entre le plasma à l'équilibre hydrostatique et des sous-groupes de gaz froids et denses en déplacement (Vikhlinin, 2001, Markevitch, 2001, Churazov, 1999, Churazov, 2003). Dans les cas où la géométrie est favorable, la mesure des variations de température et d'émissivité du plasma causées par les chocs permet de déduire la cinématique des composantes de gaz choqué. Le couplage des variations de température et d'émissivité du plasma permet de mesurer des variations d'entropie et de quantifier la dissipation d'énergie des chocs successifs subis par le plasma (Markevitch, 1999). Enfin, dans les amas relaxés, on détecte des écarts de métallicité entre les régions du plasma à l'équilibre et les régions centrales multiphasées perturbées par les courants de refroidissement (Fabian, 2002, Morris & Fabian, 2003, Sanders & Fabian, 2002). C'est dans le but d'étudier globalement les propriétés du plasma intra-amas dans les amas de galaxies en coalescence observés par ces nouveaux satellites, par exemple en les comparant aux résultats de simulations numériques, que nous avons cherché à cartographier de manière exhaustive et optimale la distribution des paramètres de densité, de température et d'entropie du plasma.

Les spectro-imageurs X du type des caméras EPIC placées au foyer des télescopes équipant le satellite XMM-Newton, sont des caméras CCD à comptage de photons. Pour chaque observation, ils fournissent une liste de photons ou d'"événements", chaque événement étant caractérisé par une coordonnée spatiale (k,l) et une énergie e . En extrayant de la liste de photons un échantillon d'événements associé à une région du plan du ciel, on peut construire l'estimateur $\hat{\theta}$ d'un paramètre physique θ du plasma intervenant dans le processus d'émissivité, par exemple la densité n , la température T , la métallicité Z , ou encore l'entropie spécifique s déduite de T et n . Le problème posé est de cartographier de manière optimale la distribution du paramètre $\theta(k,l)$, le paramètre étant localement estimé par $\hat{\theta}$ dans différents éléments de résolution spatiaux à déterminer. En l'absence d'a priori sur la géométrie de la source, on utilisera des éléments de résolution isotropes de taille variable. On pourra par exemple payer le plan du ciel en pixels carrés de coordonnée spatiale (k,l) et de largeur a . Si l'estimateur $\hat{\theta}$ est optimal, sa variance $\sigma_{\hat{\theta}}^2$ décroît quand la taille de l'échantillon augmente. Cette taille augmentant avec la surface des éléments de résolution spatiaux dans lesquels le paramètre est estimé, le choix du raffinement de maillage du plan du ciel est le résultat d'un compromis entre la résolution spatiale et l'intervalle de confiance associés à chaque estimation locale. Dans le cadre de l'analyse multirésolution, la transformation en ondelettes apporte une solution à ce compromis en codant l'amplitude des variations spatiales du paramètre $\theta(a,k,l)$ et de sa variance $\sigma_{\hat{\theta}}^2$ à l'échelle a par des coefficients de moyenne nulle dans l'espace spatio-fréquentiel. Dans cet espace, le signal $\theta(k,l)$ est caractérisé par de fortes variations localisées, tandis que des variations d'amplitude plus faible non localisées caractérisent sa variance $\sigma_{\hat{\theta}}^2$. En sélectionnant les coefficients dont l'amplitude dépasse un seuil de signification, on isole les éléments de résolution contribuant à restaurer les variations significatives du signal. La transformée en ondelettes de Haar permet d'implémenter une dualité entre cette approche spatio-fréquentielle et une statistique de comptage multi-résolutions dans l'espace direct. Dans le but de restaurer la distribution d'émissivité du plasma intra-amas $\theta = L_X$ à partir d'une distribution poissonnienne de photons, nous utiliserons un algorithme dans lequel cette transformée est couplée à un estimateur local du paramètre de Poisson, – la moyenne du nombre de photons –. Dans cet algorithme, spécialement adapté au comptage de photons à faible statistique (Jammal & Bijaoui, 2000), la transformée en ondelettes est seuillée à partir de la densité de probabilité des coefficients en ondelettes associés à un processus de Poisson stationnaire. Pour restaurer les distributions spatiales de température et d'entropie du plasma intra-amas, nous couplerons aussi la transformée en ondelettes de Haar avec des estimateurs locaux des paramètres θ et de leur variance σ_{θ} , mais ces estimateurs seront cette fois construits en utilisant l'information spectrale associée à la statistique

de photons.

La première partie de ce document est consacrée aux amas de galaxies, à leur formation et leur évolution. Après en avoir rappelé la constitution, –galaxies, plasma intra-amas, matière noire–, nous détaillerons les deux processus majeurs d’émission du plasma intra-amas ionisé: le rayonnement de freinage et les raies de désexcitation et de recombinaison des différents éléments. Sous l’hypothèse d’équilibre hydrostatique, nous évoquerons les propriétés des amas relaxés, dans lesquels on peut estimer la masse de gaz et la masse totale à partir de l’émissivité et de la température du plasma, qui sont des observables. Nous évoquerons ensuite la formation et l’évolution des amas de galaxies dans le cadre de la formation des structures et nous montrerons que les distributions de température, d’émissivité et d’entropie du plasma sont des observables privilégiées pour étudier la dynamique des amas en coalescence à travers les processus de chocs et de mélange intervenant dans le gaz.

La seconde partie est consacrée à l’imagerie et à la spectro-imagerie du plasma intra-amas. Nous montrerons comment, à travers une projection dans l’espace spatio-fréquentiel, la transformée en ondelettes permet d’isoler les variations localisées d’un signal bruité de la contribution uniforme du bruit. Dans le but de restaurer la distribution d’émissivité du plasma intra-amas à partir d’une distribution poissonnienne de photons à faible statistique, nous comparerons différents algorithmes adaptatifs de débruitage. Parmi eux, nous choisirons d’utiliser un algorithme de seuillage de la transformée en ondelettes de Haar spécialement adapté à l’estimation locale d’un paramètre de Poisson. Nous présenterons ensuite un algorithme de spectro-imagerie destiné à cartographier la température et l’entropie du plasma intra-amas, et nous discuterons de ses performances tant en imagerie qu’en spectroscopie en présentant des tests sur des observations simulées.

La troisième partie est consacrée à la présentation de quelques résultats astrophysiques obtenus en appliquant ces algorithmes d’imagerie et de spectro-imagerie à des observations récentes d’amas et de groupes de galaxies avec les spectro-imageurs EPIC équipant le foyer du télescope XMM-Newton. En imagerie, nous discuterons de la mise en évidence de l’émission étendue du plasma intergalactique dans un groupe compact de sept galaxies spirales (HCG 16). En spectro-imagerie, nous présenterons des cartes de température et d’entropie du plasma intra-amas dans deux amas de galaxies bimodaux en cours de coalescence (Abell 1750 et Abell 3921). Ces cartes mettent en évidence des zones de compression chaudes du gaz à l’interface des sous-groupes en coalescence. Enfin nous présenterons la carte en température très perturbée d’un amas de galaxies chaud et massif ayant subi des processus de coalescence et n’étant pas encore relaxé: Abell 2163.

Première partie

Formation et évolution des amas
de galaxies.

Chapitre 1

Les amas de galaxies et l'astronomie X.

1.1 Les amas et les groupes de galaxies.

Si la distribution des galaxies est considérée comme homogène et isotrope à grande échelle, elle s'est par contre structurée sous l'influence de la gravitation à des échelles inférieures à 100 Mpc, au cours de l'évolution de l'Univers. La répartition des galaxies laisse alors apparaître des surdensités filamentaires séparant des régions vides. Ces surdensités elles-mêmes ne sont pas homogènes, les galaxies ayant tendance à s'agréger le long des filaments en petites unités de quelques galaxies à quelques dizaines de galaxies, appelées groupes de galaxies. Les amas de galaxies sont situés à l'intersection des filaments. Constitués de plusieurs centaines voire plusieurs milliers de galaxies regroupées dans des volumes de 1 à 10 Mpc³, ils forment les systèmes autogravitants les plus massifs observés dans l'Univers.

Les amas de galaxies ont été intensivement étudiés en astronomie optique. G. Abell (1958) puis F. Zwicky (1961-1968) ont constitué des catalogues d'amas qui servent toujours de référence en identifiant les amas par des critères statistiques exhaustifs. Les amas de galaxies se distinguent par leur richesse, —une fonction du nombre de galaxies— et par leur régularité. Les amas réguliers sont des amas riches dont la répartition de galaxies est dense dans la région centrale et proche de la symétrie sphérique. Ils sont essentiellement constitués de galaxies elliptiques ou lenticulaires, de type E ou S0. L'amas de Coma est un amas régulier. Les amas irréguliers ne présentent pas de symétrie évidente ni de surdensité unique au centre, mais peuvent présenter des morphologies variées. Ils sont trois à dix fois moins denses que les amas réguliers et sont constitués de galaxies de tout type dont un pourcentage conséquent de spirales. L'amas de Virgo est un amas irrégulier. Les amas massifs laissent clairement apparaître des propriétés de régularité et de ségrégation morphologique des galaxies par rapport aux galaxies de champ. Ces propriétés impliquent qu'ils constituent des structures ayant évolué de manière isolée sans se disperser. Comme le temps de traversée d'un amas par une galaxie est de l'ordre de grandeur de l'âge de l'amas ($t_{cr} = 10^9$ ans, Sarazin, 1986), on en déduit que les amas sont des systèmes gravitationnellement liés. Sous cette hypothèse, on peut déterminer leur masse en étudiant la dispersion de vitesse des galaxies, alors vues comme des particules-tests dans le potentiel gravitationnel de l'amas. Ces calculs montrent que la masse des amas est largement supérieure à la masse totale des étoiles et du milieu interstellaire constitutif des galaxies. En effet la masse totale des amas ainsi dérivée est typiquement de $10^{15} M_{\odot}$ pour une luminosité de $10^{13} L_{\odot}$. Les amas de galaxies sont en fait comme toutes les structures de l'Univers essentiellement constitués de matière noire dont la présence est révélée par ses effets gravitationnels.

La matière noire et la matière baryonique des galaxies ne constituent cependant pas les seules composantes de la masse totale des amas. Une troisième composante a été mise en évidence avec la découverte dans les années 1970 d'un gaz chaud (10^8 K) et peu dense ($n \simeq 10^{-3} \text{ cm}^{-3}$), émettant un rayonnement de freinage dans le domaine des X et remplissant l'espace intergalactique dans les amas. Ce gaz a d'abord été observé dans les amas les plus proches (Virgo; Coma; Perseus, Cavaliere et al., 1971), puis dans de nombreux autres amas avec le lancement du satellite Uhuru en 1972, permettant le premier relevé complet du ciel en astronomie X. Ces premières observations ont montré que:

- L'extension de cette composante gazeuse est comparable à celle de la distribution de galaxies (Kellog et al. 1972, Forman et al., 1972).
- La vitesse d'agitation des atomes du gaz, $\frac{kT_g}{\mu m_p}$, de température T_g est comparable à la vitesse de dispersion des galaxies (Felten et al., 1966). Ce résultat s'interprète naturellement en considérant que les deux composantes de matière baryonique sont gravitationnellement liées dans un même potentiel gravitationnel.
- La luminosité X (L_X) et la vitesse de dispersion des galaxies (σ_r) sont corrélées dans de nombreux amas: $L_X \propto \sigma_r^4$ (Solinger et Tucker, 1972), ce qui implique une fraction de masse de gaz constante.

- La luminosité X et la fraction de spirales dans les amas sont inversement corrélées, les galaxies interagissant d’autant plus avec le milieu intra-amas que celui-ci est dense et que les amas sont massifs.

Les progrès de l’astronomie X depuis ces découvertes, notamment en sensibilité et en résolution spatiale, permettent aujourd’hui d’étudier en détail les propriétés du gaz chaud dans les amas et groupes de galaxies. On sait maintenant que la contribution de la masse du gaz intra-amas à la masse totale est de l’ordre de 10 – 20%, elle est donc sensiblement supérieure à la contribution des galaxies. Le rapport masse-luminosité croissant avec la masse des amas (Blumenthal et al., 1984), le rapport de masse entre les contributions du plasma intra-amas et des galaxies croît de 1 à 7 des groupes de galaxies aux amas les plus massifs (David et al., 1989, Arnaud et al., 1992). Le plasma intergalactique est considéré comme la composante baryonique dominante dans l’Univers, notamment à cause de sa contribution dans les groupes de galaxies (Fukugita, 1998). L’abondance en éléments lourds dans le plasma intra-amas, de l’ordre de la moitié de l’abondance solaire, prouve que l’interaction entre les galaxies et le milieu intra-amas a joué un grand rôle dans la formation et l’évolution des amas. Pourtant l’hypothèse d’une origine purement galactique du plasma intra-amas est exclue par l’importance de sa contribution de masse à la masse baryonique dans les amas massifs. On suppose donc qu’il s’agit essentiellement de gaz primordial accrété dans le potentiel gravitationnel des amas.

En astronomie X comme en astronomie optique, on a pu classer les amas selon leur régularité, la distribution de gaz dans les amas les plus massifs étant proche de la symétrie sphérique alors que les amas les moins massifs présentent des morphologies X très perturbées. Dans les amas massifs et relaxés, l’hypothèse d’équilibre hydrostatique permet de déduire de la température du gaz la masse totale associée au potentiel gravitationnel, tandis que la masse de la composante gazeuse est déduite de l’émissivité du gaz. Sous cette même hypothèse, les amas massifs présentent des propriétés d’autosimilarité morphologique. Par ailleurs, d’autres paramètres comme la fraction de baryons étant conservés, la température, la masse et l’émissivité de ces amas sont corrélées. Ce type de mesures pratiquées de façon systématique sur des catalogues d’amas de différents âges et de différentes masses est un indicateur de l’évolution des distributions de matière noire et baryonique dans l’Univers et fournit des contraintes sur les paramètres cosmologiques.

La distribution spatiale du gaz intra-amas est un meilleur indicateur de la forme du potentiel gravitationnel que la distribution de galaxies parce qu’elle est continue. L’étude des distributions spatiales du plasma intra-amas a révélé une structuration plus complexe que celle des distributions spatiales de galaxies, permettant de définir six à sept sous-classes d’amas (Jones & Forman, 1992), parmi lesquelles des amas doubles et des amas simples. A cette classification morphologique s’ajoute une distinction entre d’une part les amas dont la distribution du plasma est centrée sur un coeur dense et une galaxie brillante en X, d’autre part les amas dont la distribution de plasma est plus étendue, ne présentant pas de galaxie centrale brillante en X (Forman & Jones, 1990). La présence d’une galaxie brillante en X au coeur d’amas présentant par ailleurs des propriétés d’irrégularité, –faible vitesse de dispersion des galaxies, fraction élevée de spirales, faible température du plasma–, a montré que l’évolution des amas ne pouvait se résumer à un scénario dans lequel les amas évoluent comme des systèmes isolés d’un état dynamique perturbé à un état virialisé. Par ailleurs, les alignements de sous-structures observés dans les amas coïncident avec l’alignement de structures à plus grande échelle (West et al., 1995). Ces observations montrent que les amas sont actuellement en train de se former en accréant des sous-groupes de matière issus des surdensités de matière à grande échelle. Elles sont en accord avec le consensus théorique et observationnel qui conduit aujourd’hui à privilégier un processus hiérarchique de formation des structures dans le cadre de la cosmologie $\Lambda - CDM$. Si le plasma intra-amas ne constitue pas un système isolé, il est en fait soumis à une compétition entre les phénomènes d’accrétion de sous-structures et de relaxation du gaz, de temps caractéristiques comparables. On observe ainsi d’une part des amas relativement massifs et relaxés et d’autre part des amas en cours de coalescence, plus éloignés de l’état d’équilibre.

Les télescopes X récents permettent d’allier résolution spatiale et spectrale et d’étudier, en

cartographiant la température et la métallicité du gaz, les processus physiques variés intervenant au sein du plasma intra-amas. L'interaction du gaz avec les galaxies est observable dans les amas les plus proches à travers des processus d'enrichissement. La forte densité du plasma au coeur des amas relaxés favorise le refroidissement du gaz et entraîne des courants de gaz en direction du coeur, -courants de refroidissements-. Les processus de coalescence d'amas génèrent des ondes de chocs qui chauffent le gaz et des processus de mélange entre sous-groupes de températures distinctes. C'est particulièrement dans le but d'étudier ces processus de coalescence que nous présenterons des techniques d'imagerie et de spectro-imagerie du plasma intra-amas.

1.2 Le plasma intra-amas.

Le plasma intra-amas est un gaz chaud ($T_g \simeq 10^8 K$) et peu dense ($n \simeq 10^{-3} cm^{-3}$) essentiellement constitué, comme la matière primordiale, d'Hydrogène et d'Helium. Ce gaz est ionisé et optiquement mince pour le rayonnement X étant données les températures et densités considérées. Par ailleurs on peut considérer que le plasma intra-amas est à la limite coronale, c'est-à-dire que:

- les temps caractéristiques de collisions sont faibles devant le temps de refroidissement du plasma de sorte que les particules suivent une distribution de Maxwell-Boltzmann de température T_g ;
- les processus d'excitation et de désexcitation collisionnelle sont plus lents que la désexcitation radiative, de sorte que les ions sont au niveau fondamental;
- les transitions stimulées sont négligeables étant donné la dilution du champ de rayonnement.

Sous ces conditions, l'émissivité du plasma est dominée par trois processus, le rayonnement de freinage des électrons par les ions, -transitions libre-libre-, la recombinaison radiative, -transitions libre-liée- et la désexcitation à deux photons des niveaux métastables dans les ions hydrogénoïdes. Par ailleurs le plasma est à l'équilibre d'ionisation.

La composante dominante du spectre d'émission du plasma intra-amas est un continuum de rayonnement de freinage libre-libre des électrons par les ions. Le rayonnement de freinage émis par unité de temps, de fréquence et de volume par les ions de charge Z est donné par:

$$\begin{aligned} \epsilon_{\nu}^{ff} &= \frac{2^5 \pi e^6}{3 m_e c^3} \left[\frac{2\pi}{3 m_e k} \right]^{\frac{1}{2}} Z^2 n_e n_i g_{ff}(Z, T_g, \nu) T_g^{-\frac{1}{2}} \times \exp\left(\frac{-h\nu}{k T_g}\right), \\ &= n_e n_i \Lambda(T) \end{aligned} \quad (1.1)$$

où n_e et n_i sont les densités électronique et ionique, m_e la masse de l'électron, T_g la température du gaz et g_{ff} le facteur de Gaunt qui tient compte de la cinématique des collisions et d'effets quantiques. Ce terme varie peu avec ν et T_g dans les longueurs d'onde X. On déduit de (1.1) que l'émissivité du plasma intra-amas est proportionnelle au carré de la densité de gaz et que la forme du continuum est dominée par le terme $\exp\left(\frac{-h\nu}{k T_g}\right)$ et chute rapidement à haute énergie.

De nombreuses raies de recombinaison des ions hydrogénoïdes associées aux éléments lourds du plasma sont cependant détectables aux longueurs d'onde X. L'ensemble de ces processus a d'abord été étudié pour le plasma de la couronne solaire par R. Mewe (1971, 1975), puis pour le plasma intra-amas par Mewe et Gronenschild (1977, 1981, 1985). Ces auteurs ont fourni à la communauté X des modèles d'émissivité de plasmas ionisés tenant compte des processus de rayonnement de freinage, de recombinaison et de désexcitation pour tous les éléments présents: H, He, C, N, O, Ne, Na, Mg, Al, Si, S, Ar, Ca, Fe, Ni. C'est le modèle de Mewe (1985) que nous utiliserons dans nos observations.

La figure 1.1 illustre le spectre d'émission d'un plasma intra-amas isotherme à différentes températures (10^7 à $4.10^8 K$) simulé par Sarazin et Bahcall (1977). Les émissivités sont calibrées pour une sphère de 0.5 Mpc de diamètre et pour une densité de $n = 10^{-3} cm^{-3}$. On constate que le continuum de rayonnement de freinage domine la forme du spectre, mais que l'influence relative

des raies d'émission augmente au fur et à mesure que la température du plasma et donc son degré d'ionisation diminue. L'émissivité des amas froids et des groupes de galaxies peu massifs est ainsi beaucoup plus sensible aux raies d'émission que celle des amas chauds et massifs. Par ailleurs, le continuum de rayonnement de freinage étant dû aux atomes d'hydrogène alors que les raies d'émission sont dues aux éléments lourds, le rapport d'émissivité entre raies et continuum permet d'estimer la métallicité du milieu.

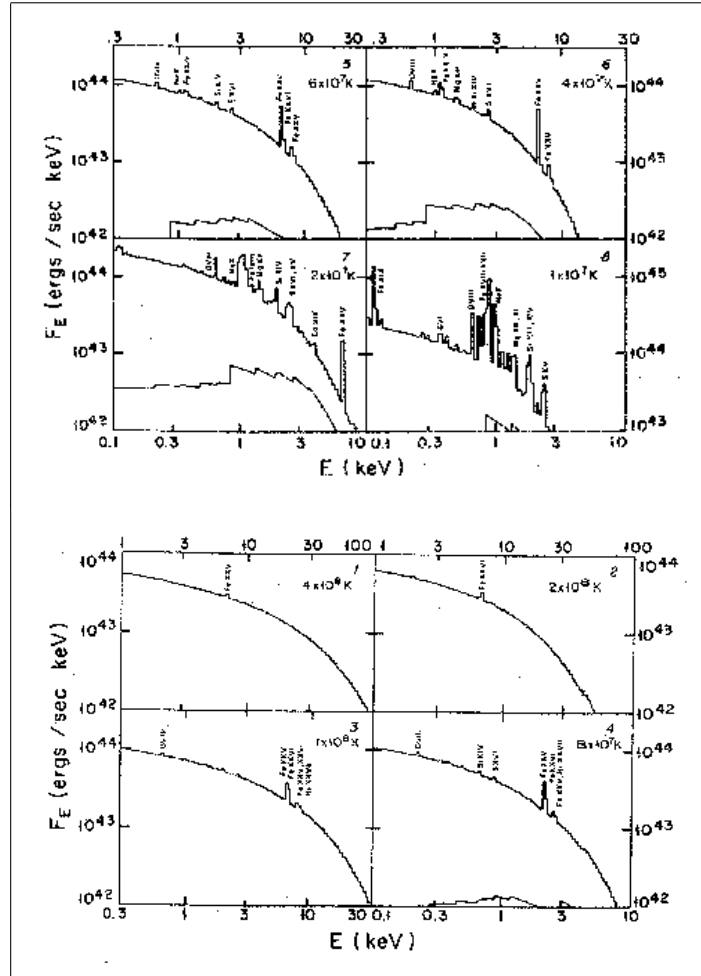


FIG. 1.1 – Spectre d'émission du plasma intra-amas à différentes températures, de 10^7 à $4 \cdot 10^8$ K, correspondant à des énergies de 1.3 à 50 keV, d'après Sarazin et Bahcall (1977).

Chapitre 2

Les amas relaxés

Le plasma intra-amas constitue en première approximation un gaz de particules autogravitantes à l'équilibre hydrostatique, qui a été accrété dans le potentiel gravitationnel de l'amas. L'hypothèse d'équilibre hydrostatique se justifie en considérant les amas de galaxies comme des systèmes isolés et relaxés, c'est-à-dire dont les surdensités se sont homogénéisées au cours de l'évolution de l'amas. Le temps de traversée de l'amas par une onde sonore est en effet beaucoup plus court que l'âge de l'amas. L'hypothèse d'autogravitation se justifie quant à elle en considérant que les vitesses d'agitation des particules de gaz $\frac{kT_g}{\mu m_p}$ impliquent un temps de traversée de l'amas par les particules de l'ordre de grandeur de l'âge de l'amas. Le plasma intra-amas se disperserait donc s'il n'était pas autogravitant (Sarazin, 1985). L'isotropie du problème conduit à exprimer l'équilibre hydrostatique dans le cadre d'un modèle sphérique, le plasma intra-amas étant caractérisé par ses profils de densité $\rho(r)$ et de température électronique $T(r)$. Le plasma intra-amas étant autogravitant, l'énergie cinétique des particules de gaz s'oppose au potentiel gravitationnel de l'amas. La température électronique T_X du plasma est donc un estimateur de la masse totale de l'amas. L'émissivité du plasma L_X est quant à elle dépendante de la densité de gaz et constitue un estimateur de la masse de gaz. Dans le cadre du modèle hiérarchique de formation des structures, les amas de galaxies se sont formés par contraction gravitationnelle au sein de surdensités de matière noire devenues instables. Ils ont alors atteint un état d'équilibre virialisé. Les dispersions de vitesse et les fractions de masse des différentes composantes de matière étant initialement homogènes, ces quantités se sont conservées et on observe aujourd'hui que les amas relaxés présentent des propriétés autosimilaires quelle que soit leur masse: mêmes profils de densité et de température, fraction de baryons constante, corrélation masse température $M_X - T_X$ et luminosité-température, $L_X - T_X$.

2.1 Profils de densité.

2.1.1 Densité de matière baryonique.

Des observations d'amas massifs et relaxés ont montré que les distributions de densité du plasma intra-amas d'une part et des galaxies d'autre part sont en accord avec un modèle où chacune des deux composantes de matière baryonique se comporterait comme une distribution de particules isotherme et autogravitante dans le potentiel gravitationnel de l'amas (Lea, 1975, Cavaliere & Fusco-Femiano, 1976, Bahcall & Sarazin, 1977). Ce modèle est celui de la sphère isotherme autogravitante de King (1962). Son profil de densité peut être approché par une distribution du type:

$$\rho(r) = \rho_o \left[1 + \left(\frac{r}{r_o} \right)^2 \right]^{-\frac{3}{2}}, \quad (2.1)$$

caractérisée par un rayon de coeur r_o . Si chaque composante de matière baryonique est isotherme et autogravitante, le plasma intra-amas est plus collisionnel que les galaxies, et la vitesse d'agitation des particules de gaz est plus élevée que celle des galaxies pour un même potentiel gravitationnel. Sous cette hypothèse, Cavaliere & Fusco-Femiano (1976) ont montré que si la distribution de densité $\rho(r)$ des galaxies suit un profil de King du type (2.1), la distribution de densité du plasma intra-amas suit un profil de densité du type $\rho_{gaz}(r)^\beta$, encore appelé modèle β :

$$\rho_{gaz}(r) = \rho_o \left[1 + \left(\frac{r}{r_o} \right)^2 \right]^{-\frac{3\beta}{2}}, \quad (2.2)$$

où $\beta = \frac{\sigma_g^2}{\sigma_X^2}$ est le rapport des vitesses d'agitation des galaxies et des particules de gaz. On observe qu'il est relativement constant d'un amas à l'autre. Arnaud et al. (2002) ont par exemple rapporté des mesures récentes effectuées sur un grand nombre d'amas chauds et relaxés de différentes masses et à différentes distances ($0.04 < z < 0.80$). Ils ont montré que les profils de densité de ces amas présentaient des propriétés homothétiques remarquables ajustables par un modèle- β avec $\beta = \frac{2}{3}$, du moins au-delà du rayon de coeur. Cette autosimilarité morphologique leur a permis

d'utiliser la distance angulaire des amas comme un indicateur de distance et de contraindre les paramètres cosmologiques.

2.1.2 Densité de matière noire

Les distributions de densité du plasma intra-amas et des galaxies étant proches des profils de densité de sphères isothermes autogravitantes, on peut supposer que la matière noire qui s'est effondrée dans les mêmes conditions que la matière baryonique non collisionnelle suit la même distribution de densité que les galaxies. On verra dans la prochaine section que cette hypothèse permet de déduire la masse totale des amas de l'ajustement des paramètres β et r_o de la distribution de densité du plasma intra-amas. La comparaison des masses totales d'amas ainsi mesurées avec d'autres estimateurs de masse, – dispersion de vitesse des galaxies, lentilles gravitationnelles –, montre a posteriori que cette hypothèse de départ était raisonnable et que la dispersion de vitesse des particules de matière noire σ_{DM} est proche de celle des galaxies.

Si le modèle des sphères isothermes autogravitantes fournit des résultats satisfaisants en ce qui concerne l'allure des profils de densité observés et les mesures de masse, il n'en reste pas moins un modèle idéalisé. Des simulations hydrodynamiques récentes de formation des structures ont permis d'affiner l'allure des profils de densité des amas. Ces simulations prennent en compte l'allure du spectre initial des fluctuations de densité (cf. section 3.2), l'évolution dynamique de l'Univers, les processus collisionnels de chocs et de turbulence dans le gaz et le couplage entre les différentes composantes de matière. Par contre, les hypothèses de symétrie sphérique, d'isothermalité et d'équilibre hydrostatique sont levées. Les simulations de Navarro, Frenk & White (1995) ont montré que les profils de densité de gaz dans les amas suivaient bien des lois autosimilaires proches des lois du type modèle- β , mais que le moment angulaire de la matière non collisionnelle était en partie transféré à la matière collisionnelle, de sorte que les profils de densité de matière noire étaient plus piqués qu'un modèle $-\beta$. Ils s'ajustent par une loi singulière en $\frac{1}{r}$ au coeur des amas:

$$\rho_{NFW}(r) \propto \frac{1}{r(1 + \frac{r}{r_s})^2}. \quad (2.3)$$

Ce profil de densité de matière noire est universel sur des échelles s'étendant des halos de matière noire galactique aux amas de galaxies (Navarro, Frenk & White, 1996).

2.2 Calculs de masses.

2.2.1 Masse de gaz.

Les hypothèses d'équilibre hydrostatique et de symétrie sphérique permettent de déduire la masse de la composante gazeuse de son émissivité et de sa température. L'équilibre hydrostatique à la distance r du centre de l'amas implique que le gradient de pression du gaz ∇P_X de densité ρ_g , s'oppose au gradient du potentiel gravitationnel de l'amas, $\nabla\Phi(r)$:

$$\nabla P_X = -\rho_g \nabla\Phi(r). \quad (2.4)$$

Soit en exprimant le potentiel gravitationnel de l'amas en fonction de sa masse totale M_t :

$$\frac{dP_X}{dr} = -\frac{GM_t\rho_X}{r^2}. \quad (2.5)$$

Par ailleurs la loi des gaz parfaits donne:

$$P_X = \frac{\rho_g k_B T_X}{\mu m_H}, \quad (2.6)$$

où m_H est la masse de l'atome d'hydrogène et μ le poids moléculaire moyen. De (2.5) et (2.6) on déduit alors l'égalité suivante:

$$\begin{aligned}
 \frac{d \ln(P_X)}{dr} &= \frac{1}{P_X} \frac{\partial P_X}{\partial r} \\
 &= \frac{\mu m_H}{\rho_G k_B T_X} \left[-\frac{GM_T \rho_X}{r^2} \right].
 \end{aligned} \tag{2.7}$$

Cette relation permet d'exprimer la masse de gaz contenue dans la sphère de rayon r en fonction des gradients logarithmiques de température et de densité, qui sont des observables:

$$\begin{aligned}
 M_t(r) &= -\frac{k_B T_X(r) r^2}{G \mu m_H} \left[\frac{d \ln(P_X)}{dr} \right] \\
 &= -\frac{k_B T_X(r) r}{G \mu m_H} \left[\frac{d \ln(P_X)}{d \ln r} \right] \\
 &= -\frac{k_B T_X(r) r}{G \mu m_H} \left[\frac{d \ln(\rho_X)}{d \ln r} + \frac{d \ln(T_X)}{d \ln r} \right].
 \end{aligned} \tag{2.8}$$

Dans le cas où le plasma est isotherme, le deuxième terme est nul et la masse de gaz s'exprime simplement par:

$$\begin{aligned}
 M_t(r) &= -\frac{k_B T_X r}{G \mu m_H} \left[\frac{d \ln(\rho_X)}{d \ln r} \right] \\
 &= -\frac{\sigma_X^2 r}{G} \left[\frac{d \ln(\rho_X)}{d \ln r} \right],
 \end{aligned} \tag{2.9}$$

où $\sigma_X^2 = \frac{k_B T_X}{\mu m_H}$ est la dispersion de vitesse des particules de gaz.

2.2.2 Masse de matière noire.

Si l'on suppose que les distributions de matière noire et de particules de gaz sont des sphères isothermes autogravitantes, l'équation (2.9) est vérifiée pour la matière noire comme pour la matière baryonique et les dispersions de vitesse des deux espèces de matière, σ_X^2 et σ_{DM}^2 , sont liées par la relation:

$$\frac{d \ln(\rho_X)}{d \ln r} = \frac{\sigma_{DM}^2}{\sigma_X^2} \frac{d \ln(\rho_{DM})}{d \ln r}. \tag{2.10}$$

Par ailleurs si la distribution de matière noire est celle d'une sphère isotherme autogravitante, son profil de densité peut être approché par un profil de King de rayon de coeur r_o . De (2.10) et (2.1) on déduit que le profil de densité de la distribution de gaz est un modèle- β paramétré par le rapport des dispersion de vitesses de matière noire et baryonique $\beta = \frac{\sigma_{DM}^2}{\sigma_X^2}$ (cf. équation 2.2). Sous les hypothèses d'isothermalité et d'équilibre hydrostatique des composantes de matière noire et baryonique, cette forme analytique de paramètres r_o et β peut être ajustée aux profils de densité observés pour déduire simultanément d'après (2.9) la masse de gaz et la masse totale de leurs profils de densités respectifs. La masse totale de l'amas ainsi déduite s'exprime par:

$$M_t(r) = 1.13 \times 10^{15} \beta \left(\frac{T}{10 \text{ keV}} \right) \times \left(\frac{r}{\text{Mpc}} \right) \frac{\left(\frac{r}{r_o} \right)^2}{\left[1 + \left(\frac{r}{r_o} \right)^2 \right]} h^{-1} M_\odot. \tag{2.11}$$

2.3 Relations autosimilaires.

La nature homothétique des profils de densité du plasma intra-amas observée dans les amas relaxés conduit à étudier les propriétés autosimilaires des amas à l'intérieur d'une sphère de rayon r_c et de densité moyenne constante, $\bar{\rho} = \delta_c \rho_c$, où δ_c est la surdensité de la sphère par rapport à la densité critique. La surdensité $\delta_c = 200$ est souvent utilisée parce que les simulations numériques montrent qu'elle correspond environ à la taille des structures virialisées. Le rayon r_{200} associé à cette surdensité est appelé "rayon viriel". Après avoir introduit les paramètres du problème et calculé l'émissivité X du plasma contenu dans une sphère de densité moyenne constante, nous montrerons que sous les hypothèses de l'équilibre hydrostatique et d'une fraction de baryons constante, la luminosité et la température du plasma intra-amas contenu dans une sphère de surdensité moyenne constante sont corrélées.

2.3.1 Emissivité d'une sphère de densité moyenne constante.

Le profil de densité $\rho(r)$ du plasma intra-amas dans la sphère de densité moyenne $\bar{\rho}$ suit une loi autosimilaire ρ_o telle que $\rho(r) = \rho_o(\frac{r}{r_c})$. Introduisons la variable invariante d'échelle $y = \frac{r}{r_c}$ et exprimons la loi autosimilaire $\rho_o(y)$ en fonction d'une fonction de densité $h(y)$ et de la densité critique:

$$\rho_o(y) = \bar{\rho} h(y). \quad (2.12)$$

La définition du rayon r_c de la sphère de densité moyenne $\bar{\rho}$ impose la condition de normalisation $\int_0^1 h(y) y^2 dy = \frac{1}{3}$. En effet, la masse de la sphère doit vérifier:

$$\begin{aligned} M(r_c) = \frac{4}{3} \pi r_c^3 \bar{\rho} &= \int_0^{r_c} \rho(r) \times 4\pi r^2 dr \\ &= \bar{\rho} r_c^3 \times \int_0^1 h(y) 4\pi y^2 dy. \end{aligned} \quad (2.13)$$

L'émissivité de la sphère de densité moyenne constante L_X , s'exprime en fonction de l'émissivité par unité de volume du plasma $\Lambda(T)$ et de la densité de matière $\rho(r)$ en introduisant la fraction de gaz $f_{gaz} = \frac{\rho_{gaz}(r)}{\rho(r)}$:

$$\begin{aligned} L_X &= \Lambda(T) \int_0^1 \rho_{gaz}^2(r) \times 4\pi r^2 dr \\ &= \Lambda(T) f_{gaz}^2 \bar{\rho}^2 r_c^3 \times \int_0^1 h^2(y) 4\pi y^2 dy \\ &= \Lambda(T) M(r_c) \times f_{gaz}^2 \bar{\rho} \times [C]. \end{aligned} \quad (2.14)$$

Dans cette expression le terme $[C = \int_0^1 h^2(y) 4\pi y^2 dy]$ est un facteur constant qui ne dépend que de la forme de la fonction de densité autosimilaire $h(y)$. Si l'émissivité du plasma est dominée par le rayonnement de freinage, on a $\Lambda(T) \propto \sqrt{T}$ et:

$$\begin{aligned} L_X &\propto \Lambda(T) M(r_c) \\ &\propto \sqrt{T} M(r_c) \end{aligned} \quad (2.15)$$

Relations masse-température et luminosité-température.

Sous les hypothèses d'équilibre hydrostatique et d'isothermalité des distributions des particules de gaz et de matière noire, la masse de la sphère de densité moyenne constante $M(r_c)$ peut être

déduite de l'expression (2.11). De (2.13) et (2.11) on déduit que $M(r_c) \propto r_c^3 \propto Tr_c$. Le rayon de la sphère r_c et la masse contenue dans la sphère de surdensité moyenne constante suivent donc des relations autosimilaires avec la température:

$$\begin{aligned} r_c &\propto \sqrt{T} \\ M(r_c) &\propto T^{\frac{3}{2}}. \end{aligned} \quad (2.16)$$

Cette relation traduit la corrélation entre la dispersion de vitesse des particules de gaz virialisé et la profondeur du puits de potentiel gravitationnel. En supposant que la fraction de baryons se conserve, elle implique aussi une corrélation entre la masse et la température du plasma intra-amas, ou encore entre sa luminosité et sa température, qui sont des observables. De (2.14) et (2.16) on déduit en effet que:

$$L_X \propto T_X^2. \quad (2.17)$$

Cette corrélation est vérifiée de manière remarquable avec des amas relaxés suffisamment chauds pour que l'émissivité du plasma soit dominée par le rayonnement de freinage et dans lesquels les courants de refroidissements sont peu prononcés. La figure (2.1) illustre la corrélation pour un relevé récent d'amas sélectionnés selon ces critères. Il existe donc un accord entre modèle analytique, simulations numériques et observations pour les amas relaxés à l'équilibre hydrostatique.

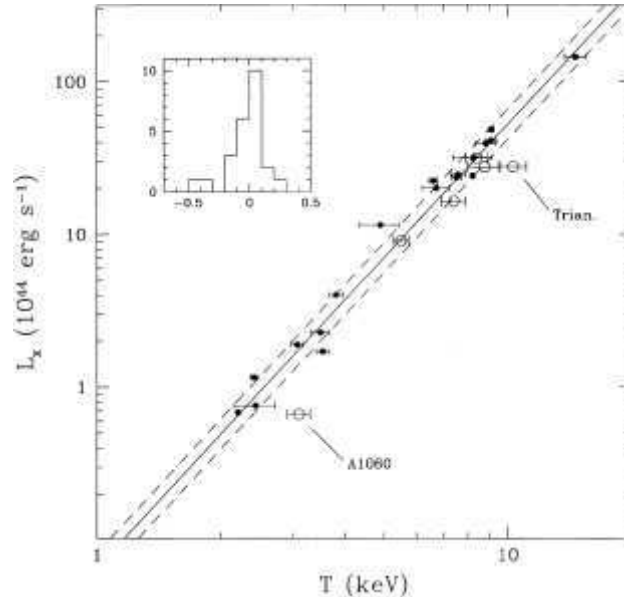


FIG. 2.1 – Corrélation luminosité-température pour un échantillon de 24 amas, d'après Arnaud & Evrard, 1999.

Chapitre 3

Formation des structures de l'Univers.

3.1 Evolution dynamique de l'Univers

Le contenu matériel de l'Univers est observable sous la forme de galaxies regroupées en filaments, groupes et amas. Il est caractérisé par une structuration hiérarchique, mais aussi homogène et isotrope à grande échelle. La composante de matière visible, grâce à des processus de rayonnement variés à diverses longueurs d'onde, ne peut expliquer à elle seule la cohésion gravitationnelle des structures. C'est pourquoi on suppose que l'essentiel de la masse des structures est constituée de halos de matière "noire", non collisionnelle. L'Univers est par ailleurs rempli d'un rayonnement homogène et thermalisé issu de l'Univers primordial. Enfin on observe que les structures constitutives de l'Univers s'éloignent d'autant plus vite les unes des autres qu'elles sont distantes. Cette propriété d'expansion d'un Univers homogène et isotrope se caractérise par la loi de Hubble qui traduit l'évolution d'un paramètre d'échelle $a(t)$, proportionnel aux distances intergalactiques:

$$\left[\frac{\dot{a}}{a} \right]_t = H(t), \quad (3.1)$$

où $H(t)$ le paramètre de Hubble à l'instant t . On modélisera finalement l'Univers comme un fluide en expansion, caractérisé par son paramètre d'échelle $a(t)$, une densité de matière, – essentiellement noire –, ρ_M , et une densité d'énergie de rayonnement ρ_R .

Dans le cadre de la théorie de la relativité générale, la courbure de l'Univers est modifiée par la densité de matière-énergie de ses principaux constituants, qui agit comme une source gravitationnelle. Cette densité ρ_T est constituée des densités d'énergie de la matière et du rayonnement, ρ_M et ρ_R , auxquelles on a ajouté le terme ρ_Λ pour rendre compte de l'influence de la constante cosmologique Λ , apparaissant dans les équations d'Einstein:

$$\rho_T = \rho_M + \rho_R + \rho_\Lambda. \quad (3.2)$$

Il est d'usage de normaliser l'ensemble de ces termes de densité à la densité critique $\rho_c = \frac{3H_0^2}{8\pi G}$, qui constitue dimensionnellement une unité "naturelle" d'énergie, et de définir les paramètres: $\Omega_i = \frac{\rho_i}{\rho_c}$. L'équation (3.1) se réécrit alors:

$$\Omega_T = \Omega_M + \Omega_R + \Omega_\Lambda. \quad (3.3)$$

La figure 3.2 illustre le comportement des trois termes de densité en fonction de la température de l'Univers, -ou du temps-. La conservation de la quantité de matière implique que la densité de matière décroît en $a^{-3}(t)$, tandis que cet effet, qui s'ajoute à la perte d'énergie des photons au cours de l'expansion -effet Doppler-, conduit à une décroissance en $a^{-4}(t)$ de la densité d'énergie de rayonnement. Enfin la "densité" associée à la constante cosmologique est un terme constant. On constate qu'au cours de son évolution, la densité d'énergie de l'Univers a successivement été dominée par l'une des trois composantes, le rayonnement dans l'Univers primordial, puis la matière pendant une grande partie du temps, puis récemment la constante cosmologique. Selon le modèle cosmologique actuellement favorisé par les observations, les valeurs des constantes Ω_M et Ω_Λ sont respectivement de l'ordre de 0,3 et 0,7, conduisant à une domination de la constante cosmologique pour un paramètre d'échelle supérieur à 0,75 fois sa valeur actuelle $a_o \left(\frac{a}{a_o} < \left[\frac{\Omega_M}{\Omega_\Lambda} \right]^{\frac{1}{3}} \simeq 0.75 \right)$.

Connaissant l'évolution de la densité d'énergie de l'Univers, on peut déduire l'évolution du paramètre d'échelle des équations d'Einstein et l'interpréter, à l'aide de l'équation de Friedmann, comme une compétition entre l'expansion et l'influence gravitationnelle des différents termes de densité:

$$\left[\frac{\dot{a}}{a} \right]^2 = H_o^2 \left[\Omega_R \hat{a}^{-3} + \Omega_M \hat{a}^{-4} + \Omega_\Lambda + (1 - \Omega_T) \hat{a}^{-2} \right], \quad (3.4)$$

où on a introduit les paramètres d'échelle $\hat{a} = \frac{a(t)}{a(t_o)}$, normalisés à leurs valeurs à l'instant présent, t_o . En multipliant (3.4) par a^2 puis en différentiant, on peut exprimer l'accélération du paramètre d'échelle:

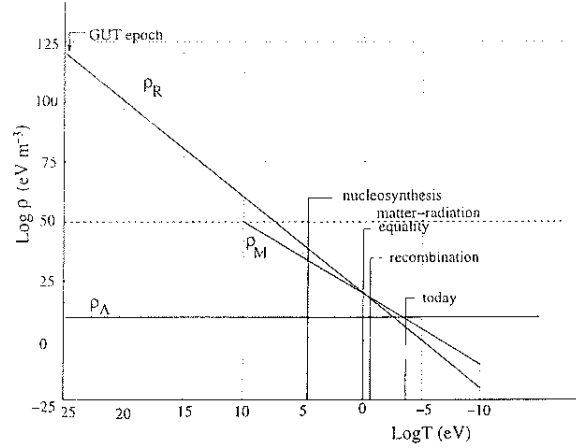


FIG. 3.1 – Densité d'énergie de radiation, de matière et densité d'énergie associée à la constante cosmologique en fonction de la température de l'Univers. D'après J. Rich.

$$\begin{aligned}\dot{a}^2 &= H_o^2 a_o^2 [\Omega_M a^{-1} + \Omega_R a^{-2} + \Omega_\Lambda a^2 + (1 - \Omega_T)] \\ \dot{a} &= -H_o^2 a_o \left[\frac{\Omega_M}{2} \hat{a}^{-2} + \Omega_R \hat{a}^{-3} - \Omega_\Lambda \hat{a} \right]\end{aligned}\quad (3.5)$$

On constate alors que les densités d'énergie de matière et de rayonnement ont pour effet de ralentir l'expansion, alors que la densité d'énergie associée à la constante cosmologique a pour effet de l'accélérer.

Durant l'essentiel du temps d'évolution de l'Univers, la densité de matière est le terme de densité dominant, on a alors $\Omega_T \simeq \Omega_M$. Par ailleurs les termes $(\Omega_R \hat{a}^{-3})$ et (Ω_Λ) sont négligeables devant le terme $(\Omega_M \hat{a}^{-4})$ et l'équation de Friedmann se ramène à la forme simplifiée suivante:

$$\left[\frac{\dot{a}}{a} \right]^2 = H_o^2 [\Omega_M \hat{a}^{-4} + (1 - \Omega_T) \hat{a}^{-2}]. \quad (3.6)$$

On distingue alors deux possibilités d'évolution du paramètre d'échelle:

- Si $\Omega_M = \Omega_T \leq 1$, \dot{a} est toujours positif et l'Univers, qualifié d'Univers ouvert, est en expansion éternelle.
- Si $\Omega_M = \Omega_T > 1$, \dot{a} est positif tant que le paramètre d'échelle est inférieur à $a_{max} = a_o \frac{\Omega_M}{\Omega_T - 1}$, s'annule si $a = a_o$, est négatif sinon. L'Univers, alors qualifié d'Univers fermé, est en expansion jusqu'à a_{max} , puis se contracte.

Les observations favorisent une valeur de Ω_T proche de 1, c'est à dire un Univers proche de la densité critique pendant la période de domination de la matière.

Enfin durant la période de domination de la constante cosmologique, la plus récente, c'est le terme (Ω_Λ) qui domine les termes $(\Omega_R \hat{a}^{-3})$ et (Ω_M) , et le paramètre d'échelle croît exponentiellement: $a(t) \propto \exp(H_o \Omega_\Lambda^{\frac{1}{2}} t)$.

3.2 Formation des structures

L'univers observable, de densité moyenne $\bar{\rho}$, est aujourd'hui structuré de manière hiérarchique avec de fortes surdensités de matière à différentes échelles: $\frac{\delta \rho}{\bar{\rho}} \Big|_{t_o} > 1$. Or l'analyse du rayonnement de fond cosmologique montre que l'univers était par le passé beaucoup plus homogène. Ce rayonnement a été émis au temps de la recombinaison t_{rec} , quand la température de l'Univers est

devenue suffisamment basse pour que la matière se découple du rayonnement. C'est un rayonnement thermalisé qui ne laisse apparaître que de très faibles inhomogénéités dans sa distribution de température et donc dans la distribution de matière qui l'a engendré. Sous l'hypothèse de perturbations adiabatiques, on a en effet, d'après les observations du fond diffus sur des échelles de 10 environ, $\left. \frac{\delta \rho}{\rho} \right|_{t_{rec}} \simeq \left. \frac{\delta T}{T} \right|_{t_{rec}} = 10^{-5}$. Toute la problématique de la formation des structures est de comprendre en détail comment les inhomogénéités de faible amplitude de l'Univers primordial ont conduit à la structuration hiérarchique observée aujourd'hui. La formation des structures se déroule en deux phases successives. Pendant la première phase, dite linéaire, avec $\frac{\delta \rho}{\rho} < 1$, on peut linéariser les équations de l'hydrodynamique pour étudier l'amplification des perturbations d'un Univers à l'équilibre hydrostatique. Pendant la deuxième phase, dite non-linéaire, avec $\frac{\delta \rho}{\rho} > 1$, on doit étudier séparément le comportement des structures en formation et celui du reste de l'Univers.

La phase linéaire.

L'amplification des inhomogénéités de matière $\frac{\delta \rho}{\rho}$ est due au processus d'instabilité gravitationnelle, développé par Jeans. Dans un fluide collisionnel, elle est le résultat d'une compétition entre la pression qui tend à effacer les inhomogénéités par la propagation d'ondes sonores et la gravitation qui les amplifie. Or le temps caractéristique d'effondrement gravitationnel d'une structure ($t_{eff} \simeq [G\rho]^{-\frac{1}{2}}$) est indépendant de sa masse M ou de sa taille caractéristique L , tandis que le temps de traversée de la structure par une onde sonore de vitesse c_s , $t_s \simeq c_s^{-1}L$, en dépend. Il existe donc une longueur caractéristique L_J , appelée longueur de Jeans, au-delà de laquelle la gravitation l'emporte sur la pression et la structure s'effondre:

$$\begin{aligned} L_J &= c_s t_{eff} \\ &= c_s G \rho^{-\frac{1}{2}}. \end{aligned} \quad (3.7)$$

On définit aussi une masse critique d'instabilité gravitationnelle appelée masse de Jeans, M_J et correspondant à la masse d'une structure de longueur caractéristique L_J . Enfin dans un fluide non-collisionnel, on peut appliquer le même formalisme en considérant que c'est la dispersion de vitesse des particules qui s'oppose à la gravitation pour effacer ou amplifier les inhomogénéités.

Pendant la période de domination du rayonnement, l'Univers était un gaz de particules relativistes, de pression $p = \frac{\rho}{3}$, correspondant à une vitesse du son proche de la vitesse de la lumière: $c_s = \frac{c}{\sqrt{3}}$. L'ordre de grandeur de la longueur de Jeans était donc de $L_J = c G \rho^{-\frac{1}{2}}$, correspondant au rayon de l'Univers observable $L_{horizon}$. A partir du moment où la matière a dominé la dynamique de l'Univers, $-z < z_{eq}$, le gaz est devenu non relativiste et la vitesse du son, ainsi que la longueur de Jeans ont chuté de plusieurs ordres de grandeur: $c_s \ll c$, $L_J \ll L_{horizon}$.

Sous l'hypothèse d'une matière essentiellement noire et non collisionnelle, -Cold Dark Matter-, toutes les fluctuations de longueur comprise entre les longueurs de Jeans avant et après l'équivalence matière-rayonnement commencent à croître à z_{eq} , tandis que les fluctuations de longueur supérieure à la taille de l'horizon à z_{eq} ont commencé à croître plus tôt. Il s'oppose aux scénarios de matière noire relativiste, -par exemple constituée de neutrinos massifs-, dans lesquels les fluctuations sont amorties après z_{eq} et les petites structures se forment plus tardivement par fragmentation. Le scénario de matière sombre froide est actuellement favorisé par rapport aux scénarios de matière noire relativiste, qui permettent difficilement expliquer l'amplitude de l'amplification des perturbations observée entre l'Univers primordial et l'Univers actuel.

Notons que ce scénario n'est valable que pour la matière noire qui domine la dynamique de l'Univers. En effet, au cours l'évolution de l'Univers, la matière baryonique a été successivement couplée au rayonnement puis découplée du rayonnement à la "recombinaison", $-z > z_{rec} > z_{eq}$, quand la température de l'Univers est devenue suffisamment basse pour que la densité d'énergie de rayonnement soit inférieure à l'énergie de dissociation des atomes d'hydrogène. Pendant la période de couplage matière-rayonnement, les fluctuations de densité de matière baryonique ont été amorties par la pression de radiation sur des échelles de distance correspondant à un temps de

diffusion des photons inférieur au temps caractéristique d'effondrement des structures. C'est seulement après la recombinaison que la distribution de matière baryonique a pu suivre la distribution de matière noire dont la croissance des perturbations avait déjà commencé à toutes les échelles à $z = z_{eq}$.

La phase non-linéaire.

La croissance des perturbations se poursuivant à toutes les échelles après $z = z_{eq}$, la densité d'énergie de l'Univers est alors dominée par la densité de matière et l'évolution de l'Univers suit l'équation de Friedmann simplifiée pour un univers dominé par la matière (3.6), par ailleurs sa densité est proche de la densité critique ($\Omega_M = 1$).

Le comportement d'une surdensité locale peut-être étudié qualitativement en utilisant le modèle sphérique. La figure 3.2 illustre une surdensité sphérique isolée de densité $\rho > \rho_o$, supérieure à la densité critique. L'application du théorème de Gauss montre que le potentiel gravitationnel auquel est soumis la surdensité ne dépend que de sa densité interne ρ . La surdensité peut donc être étudiée de manière isolée du reste de l'Univers. Elle constitue alors un mini-Univers pour lequel l'équation (3.6) s'applique. La densité de l'Univers ρ_o étant critique, la surdensité est sur-critique et constitue un mini-Univers fermé dont l'expansion va s'arrêter puis décroître. La surdensité va alors s'effondrer jusqu'à atteindre un état d'équilibre virialisé, isolé du reste de l'univers en expansion.

Le scénario hiérarchique.

Dans le cadre d'un scénario où la matière est essentiellement constituée de matière sombre froide, les structures aujourd'hui observées à toutes les échelles ont commencé à se former à l'équivalence matière-rayonnement, par instabilité gravitationnelle de fluctuations de densité de matière issues de l'Univers primordial. Si l'origine de ces fluctuations primordiales sort encore du cadre de la physique actuelle, la structuration de l'Univers aujourd'hui observée impose des contraintes sur la nature de leur spectre d'échelle. L'absence d'échelle caractéristique dans la fonction de corrélation des structures actuelles suggère de modéliser ce spectre par une loi de puissance:

$$\frac{\delta\rho}{\rho} = \frac{\delta M}{M} = AM^{-\alpha}, \quad (3.8)$$

une valeur de α positive assurant l'homogénéité de l'Univers à grande échelle. La valeur $\alpha = \frac{3}{2}$ s'accorde bien avec les observations. Par ailleurs, elle correspond à un spectre de fluctuations particulier, dit de "Harrison-Zeldovich", qui a deux propriétés théoriques intéressantes. D'une part, il est prévu par les modèles d'inflation, d'autre part il correspond à une amplitude des fluctuations invariante d'échelle à leur entrée dans l'horizon.

Un spectre en loi de puissance de type Harrison-Zeldovich implique que l'amplitude des perturbations décroît avec l'échelle de sorte que ce sont les plus petites perturbations qui deviennent non linéaires les premières. Dans ce scénario, appelé modèle hiérarchique, les structures de masse supérieure à la masse de Jeans à la recombinaison se forment rapidement. Cette masse de $10^6 M_\odot$ correspond à celle des amas d'étoiles. Les structures plus petites se forment par fragmentation, tandis que les structures plus grandes, -galaxies, groupes et amas de galaxies-, se formeront plus tardivement, par accréation de sous-structures. Les amas de galaxies sont probablement les plus grandes structures actuellement observables et les dernières structures formées. En effet si la densité d'énergie de l'Univers est maintenant dominée par la constante cosmologique le terme de densité de matière n'est plus dominant dans l'équation (3.4), la gravitation ne peut plus s'opposer à la croissance exponentielle du paramètre d'échelle et les structures ne peuvent plus s'effondrer.

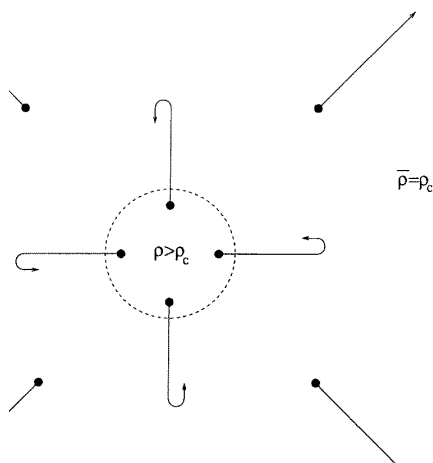


FIG. 3.2 – *Collapse gravitationnel d'une surdensité sphérique. D'après J. Rich.*

Chapitre 4

Evolution des amas de galaxies dans le cadre de la formation des structures.

La formation des structures de l'Univers démarre par une phase d'instabilité gravitationnelle et se poursuit par une phase de contraction pendant laquelle la densité de matière augmente jusqu'à atteindre un nouvel état d'équilibre virialisé (cf. section 3.2). Dans les amas de galaxies, la phase de contraction se déroule encore aujourd'hui et se manifeste par l'accrétion de sous-groupes de galaxies de masses diverses dans les amas de galaxies en coalescence. L'état d'équilibre final est celui des amas relaxés. On observe que cet état final se caractérise par une dispersion de vitesses et une distribution de densité différentes pour le plasma intra-amas d'une part, la matière noire et les galaxies d'autre part (cf. section 2). En effet, si la phase de contraction de la matière noire et des galaxies, –peu collisionnelles–, est peu dissipative, la phase de contraction du plasma intra-amas se caractérise par une relaxation violente au cours de laquelle les processus de chocs jouent un rôle important dans la dissipation de l'énergie cinétique des particules de gaz. Nous montrerons que ces processus se manifestent par de fortes variations de variables d'état du plasma intra-amas, telles que la température et l'entropie, qui sont des observables directes ou indirectes. Puis nous présenterons les résultats de simulations numériques d'amas en coalescence dans lesquelles les processus hydrodynamiques intervenant dans la phase baryonique sont couplés à la dynamique de la matière noire, dominée par la seule gravitation.

4.1 Processus de chocs adiabatiques

L'accrétion des sous-groupes de galaxies en coalescence les uns sur les autres est caractérisée par des vitesses légèrement supérieures à la vitesse du son dans le plasma intra-amas, ce qui entraîne des processus de chocs. Ces chocs jouent un rôle important dans le chauffage du plasma intra-amas, en dissipant l'énergie cinétique des sous-groupes en interaction. En traversant une onde de choc le gaz initialement supersonique devient subsonique et passe d'un état d'équilibre thermodynamique local à un autre. Cette transition se traduit notamment par des variations significatives de la vitesse d'agitation des particules, u , de la pression P et de la densité de gaz ρ . On peut déduire l'amplitude de ces variations des trois "conditions de saut" d'Hugoniot-Rankine qui traduisent la conservation la masse, du moment cinétique et de l'énergie interne du gaz, à la traversée de la zone de choc (Landau & Lifshitz, 1959):

$$\rho_2 u_2 = \rho_1 u_1, \quad (4.1)$$

$$\rho_2 u_2^2 + P_2 = \rho_1 u_1^2 + P_1, \quad (4.2)$$

$$\frac{1}{2} u_2^2 + h_2 = \frac{1}{2} u_1^2 + h_1, \quad (4.3)$$

où les indices 1 et 2 se réfèrent aux quantités physiques avant et après le passage du choc, exprimées dans le référentiel du choc et où h est l'enthalpie spécifique, soit pour un gaz parfait:

$$h = \frac{\gamma}{\gamma - 1} \frac{P}{\rho}. \quad (4.4)$$

En introduisant la vitesse du son dans le plasma, $c_s^2 = \gamma \frac{P}{\rho}$ et le nombre de Mach du choc, $M_1 = \frac{u_1}{c_s}$, on déduit de (4.1), (4.2) et (4.3) que:

$$\frac{\rho_2}{\rho_1} = \frac{u_2}{u_1} = \frac{(\gamma + 1) M_1^2}{(\gamma + 1) + (\gamma - 1)(M_1^2 - 1)}, \quad (4.5)$$

$$\frac{P_2}{P_1} = \frac{(\gamma + 1) + 2\gamma(M_1^2 - 1)}{\gamma + 1}, \quad (4.6)$$

$$\frac{T_2}{T_1} = \frac{[(\gamma + 1) + 2\gamma(M_1^2 - 1)] [(\gamma + 1) + (\gamma - 1)(M_1^2 - 1)]}{(\gamma + 1)^2 M_1^2}, \quad (4.7)$$

Comme $M_1 > 1$, on a $P_2 \geq P_1$, $\rho_2 \geq \rho_1$ et $T_2 \geq T_1$. A la traversée d'une onde de choc, le gaz ralentit –et devient subsonique–, mais sa température et sa densité, qui sont des observables,

augmentent brusquement (cf. figure 4.1). Ces deux dernières transitions se traduisent par une augmentation de l'entropie spécifique s du gaz :

$$\Delta s = s_2 - s_1 = \frac{3}{2}k \log \left[\left(\frac{T_2}{T_1} \right) \left(\frac{\rho_2}{\rho_1} \right)^{-\frac{2}{3}} \right]. \quad (4.8)$$

C'est en élevant l'entropie spécifique des particules de gaz que les processus de choc dissipent l'énergie cinétique initiale des sous-groupes de gaz en interaction et chauffent le plasma. On remarquera que si les lois de conservation (4.1), (4.2) et (4.3) sont réversibles, c'est l'application du second principe de la thermodynamique ($\Delta s > 0$) qui rend les transitions (4.5), (4.6) et (4.7) irréversibles imposant la condition $u_1 > u_2$ et en interdisant l'existence de chocs de "raréfaction", où du gaz subsonique deviendrait supersonique. L'entropie spécifique du plasma peut être mesurée indirectement à partir de la température et de la densité.

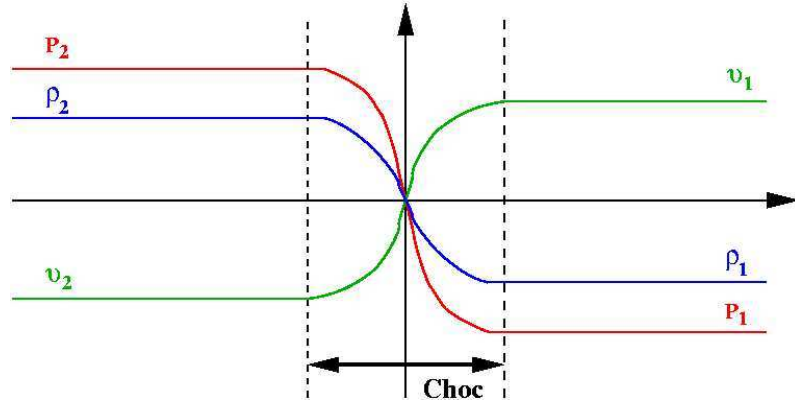


FIG. 4.1 – Modification des variables d'état du plasma à la traversée d'une onde de choc: vitesses d'agitation, v_1 et v_2 , pressions, P_1 et P_2 , densités, ρ_1 et ρ_2 .

Il existe plusieurs types de configuration dans lesquels des processus de chocs se développent dans le plasma intra-amas. Une onde de choc peut se propager librement dans le plasma intra-amas, être arrêtée au contact d'une surdensité de gaz, ou encore se propager en avant d'une surdensité de gaz en déplacement supersonique, –choc en arc–. Si une onde de choc se propage librement, elle élève la température du plasma qu'elle a traversé et on observe deux régions de température distinctes séparées par l'onde de choc en train de se propager. Si elle est arrêtée ou se propage en avant d'un coeur dense l'élévation de température causée par le choc est effacée par l'arrivée de la surdensité et on observe simplement un front chaud et dense localisé dans la zone de choc. Le chauffage du plasma causé par le passage du front de choc ne se manifeste pas alors par une élévation de la température mais de l'entropie spécifique du gaz. Contrairement à la distribution de température, la distribution spatiale de l'entropie du gaz dans le plasma garde donc la mémoire du passage des fronts de choc.

Observationnellement, les ondes de chocs ou de compression se caractérisent par une élévation simultanée de la température et de la densité du plasma ($\rho_2 > \rho_1$ et $T_2 > T_1$). Elles se distinguent des discontinuités de contact qui peuvent se développer en avant de coeurs denses en déplacement dans le plasma au repos, plus chaud ($\rho_2 < \rho_1$ et $T_2 > T_1$).

Les équations (4.5), (4.6) et (4.7) s'appliquent pour les ondes de compression comme pour les ondes de choc. On peut les inverser pour déduire les transitions de vitesses de gaz des transitions de température et de densité du plasma (Markevitch et al. 1999):

$$\Delta v_S = \left[\frac{kT_1}{\mu m_p} (C - 1) \left(\frac{T_2}{T_1} - \frac{1}{C} \right) \right]^{\frac{1}{2}}, \quad (4.9)$$

où $C = \frac{\rho_2}{\rho_1}$ est le facteur de compression. Le facteur C se déduit aussi de la discontinuité de température, –observationnellement moins sensible aux effets de projection que la discontinuité de densité– :

$$\frac{1}{C} = \left[\frac{1}{4} \left(\frac{\gamma+1}{\gamma-1} \right)^2 \left(\frac{T_2}{T_1} - 1 \right)^2 + \frac{T_2}{T_1} \right] - \frac{1}{2} \frac{\gamma+1}{\gamma-1} \left(\frac{T_2}{T_1} - 1 \right). \quad (4.10)$$

Si la configuration est favorable, –pas ou peu d'effets de superposition, coalescence dans le plan du ciel–, on peut déduire des transitions de vitesses de gaz la cinématique des chocs. Par exemple dans le cas de deux sous-groupes de masse égale en coalescence bimodale (cf. A1750, Belsole et al., 2004), une onde de compression stationnaire se développe entre les deux sous-groupes et la valeur de la vitesse relative d'un sous-groupe par rapport à l'autre est double de la valeur de la vitesse de déplacement d'un des sous-groupes, –i.e. du gaz–, par rapport à la région de choc au repos: $v_c = 2 \times \Delta v_s$. Si un coeur dense se déplace à vitesse supersonique on peut observer à la fois une discontinuité de contact autour du coeur et une onde de choc en avant du déplacement, (cf. A3667, Vikhlinin et al., 2001). La vitesse du coeur dense est alors égale à la vitesse du front de choc se déplaçant par rapport au plasma au repos: $v_c = \Delta v_s$.

4.2 Simulations numériques d'amas en coalescence.

En couplant les processus hydrodynamiques de chocs et de turbulence intervenant dans la phase baryonique avec la dynamique de la matière noire dominée par la gravitation, les simulations numériques permettent d'étudier l'évolution des variables d'état du plasma au cours des coalescences d'amas et fournissent des points de comparaison avec les observations d'amas à différents stades de leur évolution. Certaines simulations privilégient l'étude des propriétés du milieu intergalactique et modélisent les amas en coalescence comme des systèmes isolés. Elles permettent d'explorer différents scénarios de coalescence en fonction des conditions initiales, –masse, température, paramètre d'impact des sous-groupes–. D'autres simulations intègrent les amas dans la formation des structures de l'Univers à grande échelle. Ces simulations sont plus proches de la réalité et permettent de prendre en compte l'accrétion continue de sous-groupes au cours du temps, mais elles rendent les conditions initiales dépendantes du scénario global.

4.2.1 Simulation de coalescences bimodales.

La figure 4.2 illustre une simulation numérique de coalescence bimodale de deux amas isolés (Ricker & Sarazin, 2001). Les deux sous-groupes sont supposés initialement relaxés, de masses respectives M1 et M2 (cf. diagramme). La distribution de gaz suit un profil de densité du type modèle- β , tandis que la distribution de matière noire suit un profil plus piqué du type Navarro-Frenk-White (cf. section équation (2.3)). La collision des deux sous-groupes est étudiée pour différents paramètres d'impact. L'image de gauche illustre l'évolution de la distribution de température du plasma pendant une période de 4 milliards d'année incluant la collision. L'image de droite illustre l'évolution de la température moyenne et de la luminosité bolométrique L_X du plasma. L'origine de l'axe des temps est fixé au maximum d'émissivité.

Avant la collision, les deux sous-groupes sont en train de chuter l'un sur l'autre et on observe le développement d'un front de choc en arc autour du sous-groupe moins massif, dont le plasma est le moins dense. Les auteurs ont aussi simulé une collision symétrique de deux amas de masse égale, le front de choc est alors rectiligne. Au moment de la collision, l'émissivité bolométrique et la température moyenne du plasma sont maximales, –origine des temps–. Leurs valeurs ont respectivement quadruplé et doublé par rapport aux valeurs initiales quand les sous-groupes sont encore éloignés. Les processus de coalescence ont ainsi une forte influence sur l'amplitude de ces deux observables par rapport à l'état relaxé. Si le paramètre d'impact de la collision n'est pas nul, le front de choc en avant du sous-groupe le moins massif est arrêté par le sous-groupe massif et se tord, laissant apparaître une structure en "S". La collision provoque la rupture des fronts de chocs

et le départ d'ondes de chocs radiales. Elle ne détruit pas les deux-groupes, qui poursuivent leur chemin même si le paramètre d'impact est nul à cause de l'effet d'entraînement des coeurs denses de matière noire.

Poursuivie pendant une durée de 15 milliards d'années, cette simulation montre que les deux sous-groupes peuvent se croiser plusieurs fois avant de fusionner en un seul amas. En effet, si la matière baryonique dissipe rapidement son énergie cinétique initiale, il n'en va pas de même pour la matière noire. Ainsi les halos de matière noire oscillent l'un vers l'autre et empêchent la relaxation rapide du gaz. A la fin de la simulation les deux sous-groupes ont fusionné mais les distributions d'entropie sous encore très perturbées et ont gardé les traces des processus de chauffage engendrés par les chocs successifs. En outre le profil d'entropie du gaz est plus élevé dans la région centrale des amas que dans les amas relaxés à cause de mouvements de convection qui continuent à perturber le plasma et à ramener du gaz de forte entropie des régions externes vers le coeur.

4.2.2 Formation d'un amas massif au sein d'une simulation de formation des structures à grande échelle .

La figure 4.3 illustre la formation d'un amas massif au sein d'une simulation de formation des structures à grande échelle sur une échelle de temps de deux milliards d'années. La simulation a été menée en deux étapes à l'aide d'un algorithme à maillage adaptatif de formation de structures dans le cadre de la cosmologie Λ -CDM (Teyssier 2002). Dans une première étape, une centaine d'amas est formée dans un rayon de $400 h^{-1}$ Mpc. Dans une seconde étape un amas particulier est sélectionné et extrait dans un rayon de $6 h^{-1}$ Mpc et formé avec un maillage plus serré et un plus grand nombre de particules.

Les simulations de ce type permettent de construire un catalogue varié d'amas dans différentes phases d'évolution. La prise en compte des conditions initiales issues de la simulation à grande échelle permettent de rendre compte de la complexité du processus global de formation des structures. En effet les amas ne constituent pas des systèmes isolés mais accrètent continuellement des sous-groupes à des intervalles de temps inférieurs à leur temps caractéristique de relaxation.

Les deux colonnes de droite illustrent l'évolution des distributions de densité de la matière noire et du gaz intergalactique au cours du temps. Les deux colonnes de gauche illustrent l'évolution des distributions de température et d'émissivité du plasma. On observe la coalescence de deux structures massives simultanée à la chute d'autres sous-groupes moins massifs. La distribution de densité ou d'émissivité du plasma montre que la composante baryonique des deux structures dominantes a fusionné en un seul groupe à la fin de la simulation, tandis que les deux surdensités de matière noire se sont croisées et sont encore séparées. Comme dans la simulation de Ricker & Sarazin, les cartes de température mettent en évidence des fronts de choc en avant des sous-groupes en coalescence avant qu'ils ne fusionnent, puis le développement d'ondes de chocs radiales après la collision. La carte en température de la fin de la simulation montre une séparation franche entre les régions centrales chaudes déjà traversées par ces ondes de choc et les régions externes encore froides, non choquées. Ce scénario pourrait expliquer l'observation de forts gradients dans les profils de température d'amas non relaxés.

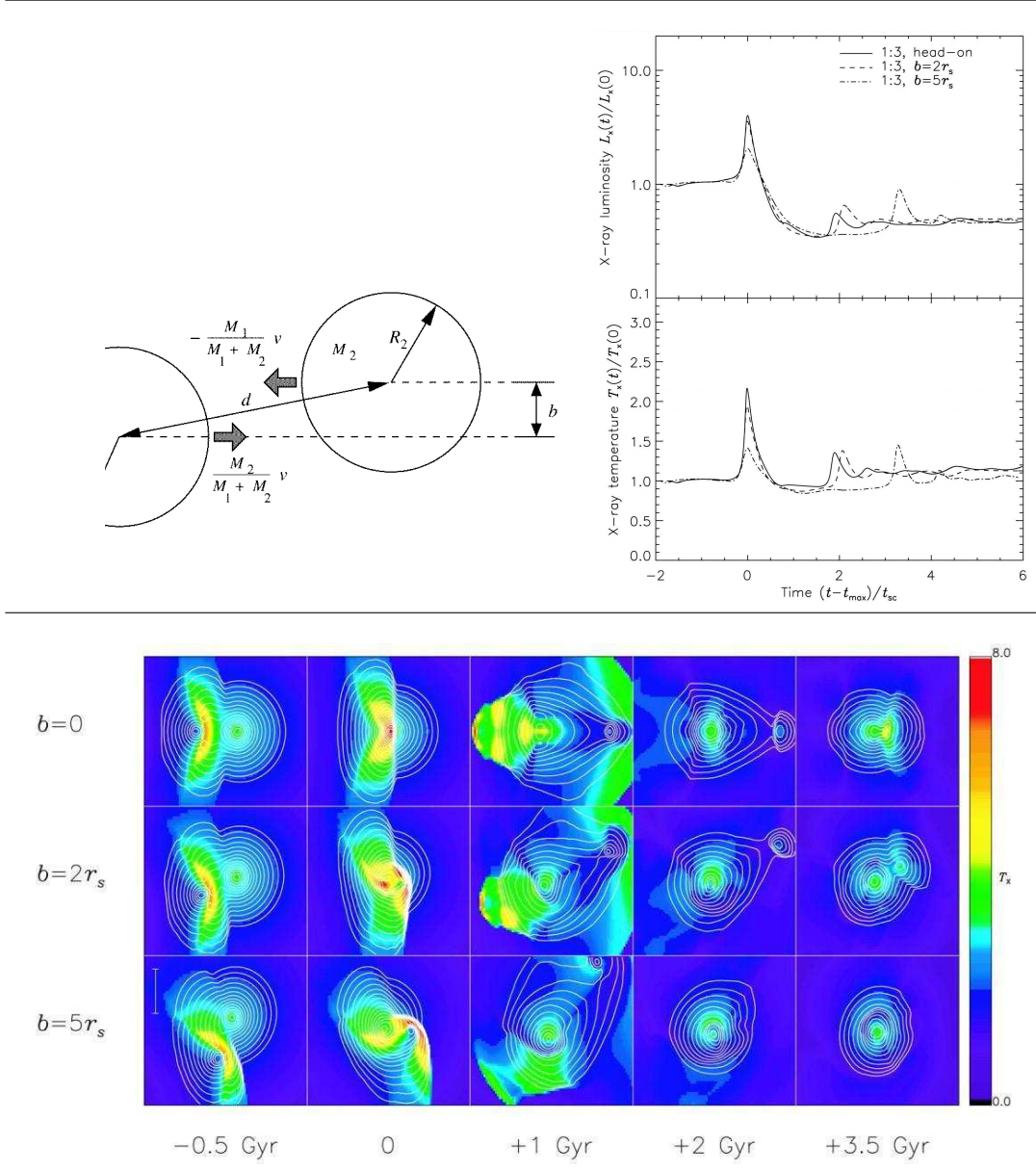


FIG. 4.2 – Simulation numérique de coalescence de deux amas relaxés, de masses respectives M_1 et M_2 et de rayons R_1 et R_2 , d'après Ricker & Sarazin 2001. La collision a pour paramètre d'impact b avec une vitesse relative v d'un amas par rapport à l'autre. Image du haut à gauche: cinématique du choc. Image du haut à droite: Evolution de l'émissivité bolométrique et de la température moyenne du plasma au cours du temps. Image du bas: Evolution de la distribution de température du plasma au cours du temps pour trois valeurs du paramètre d'impact de la collision. La température croît du bleu au rouge. La distribution d'émissivité a été superposée par des contours.

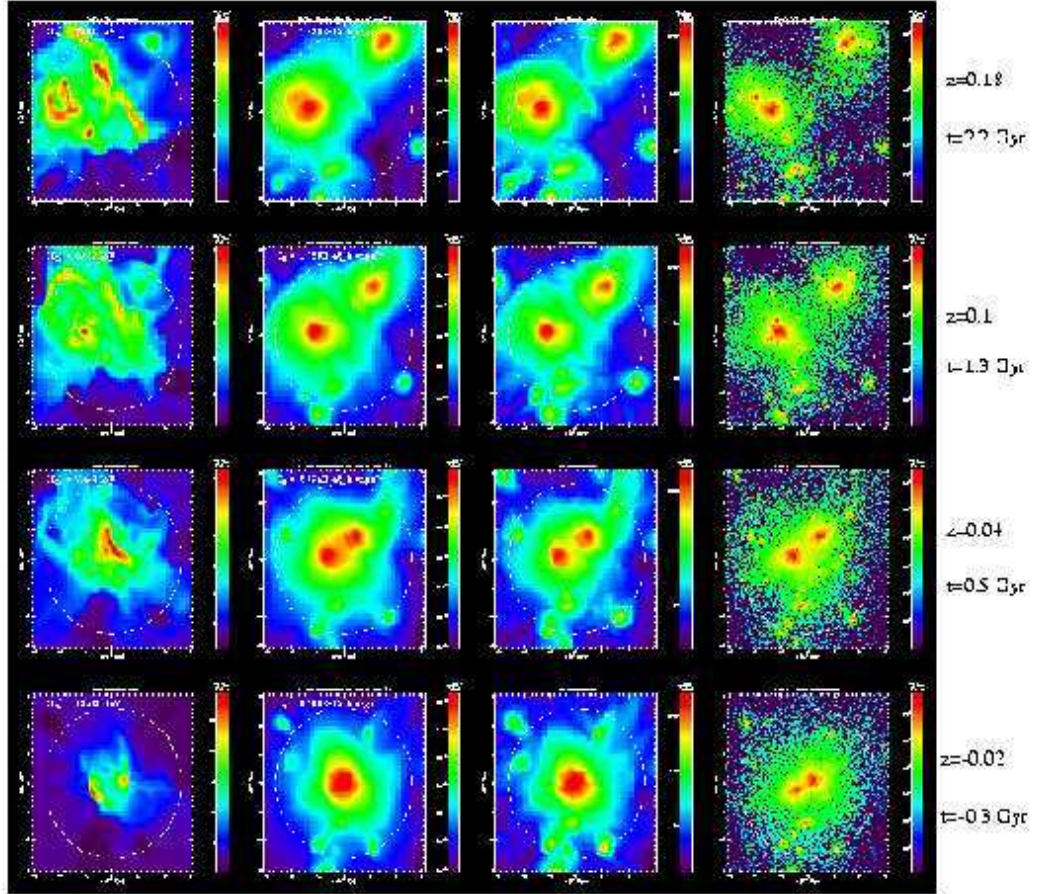


FIG. 4.3 – Simulation numérique d'amas en cours de fusion dans le cadre d'une simulation de formation des structures à grande échelle (Λ -CDM), d'après Teyssier & Gonzales, 2002. Le temps s'écoule de haut en bas. Toutes les quantités physiques sont projetées sur le plan du ciel. Colonnes de gauche à droite: (1) Distribution de température du plasma intra-amas. (2) Brilliance de surface du plasma. (3) Distribution de densité du plasma. (4) Distribution de matière noire.

Deuxième partie

Imagerie et spectro-imagerie X.

Chapitre 1

Le télescope XMM-Newton et les spectro-imageurs EPIC

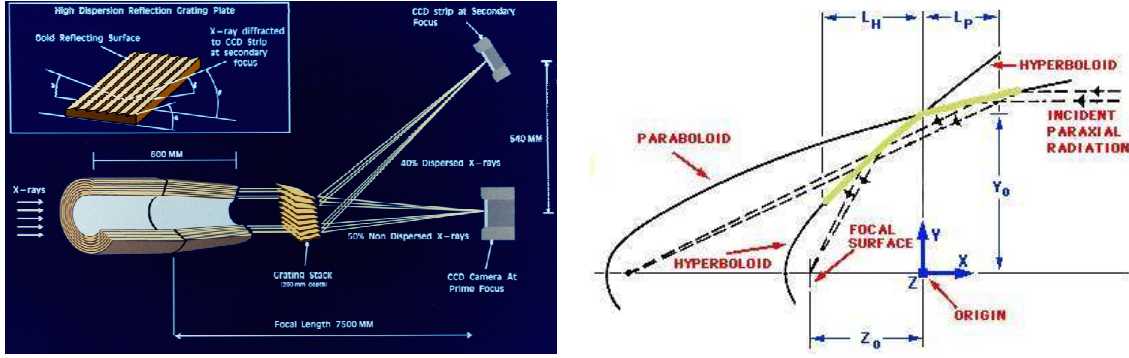


FIG. 1.1 – Droite: Chemin optique des rayons lumineux pour les deux télescopes équipés des spectro-imageurs MOS (Metal Oxide Semi-conductor) et du spectromètre RGA (Reflection Grating Spectrometer). Gauche: Chemin optique d'un rayon lumineux sur le miroir en incidence rasante.

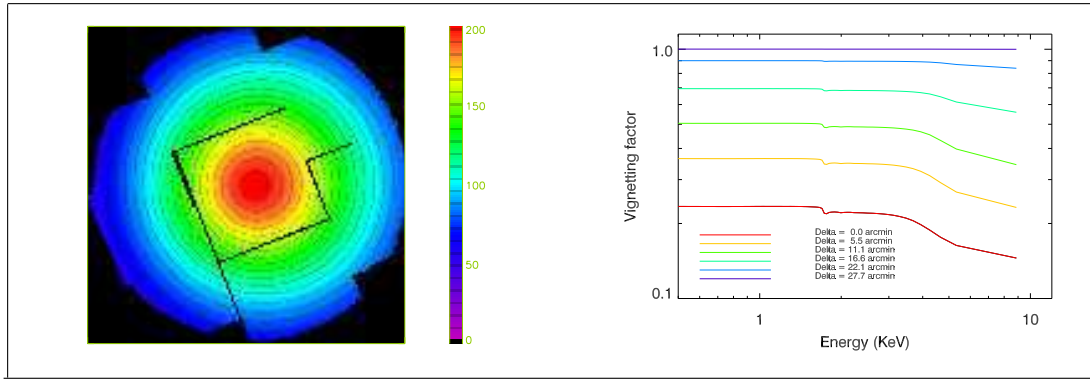


FIG. 1.2 – Aire effective de la caméra MOS1. Gauche: Aire effective totale. L'anisotropie de l'aire effective par rapport à l'axe focal est due à la dispersion par le RGA. Droite: Evolution spectrale de l'aire effective en fonction de la distance au foyer.

1.1 Présentation du télescope

Le télescope X européen XMM-Newton (X-ray Multi-mirror Mission) a été lancé en décembre 1999. Il parcourt une orbite elliptique d'une période de 48 heures dont la Terre est un des deux foyers. Le rayonnement des ceintures de radiation empêchant toute observation à une élévation inférieure à 40000 km, seules 40 heures effectives d'observation (145 ks) sont possibles au cours d'une période. Le satellite XMM-Newton est en fait composé de trois télescopes X dont les objectifs sont composés de 58 miroirs concentriques recouverts d'or et focalisant les rayons lumineux en incidence rasante. La distance focale de chaque télescope est de 7.5 m et son aire effective géométrique de 1500 cm^{-2} . Leur résolution spatiale se caractérise par une fonction d'étalement au foyer de 6 secondes d'arc de largeur à mi-hauteur. La réflexion en incidence rasante sur les miroirs des télescopes entraîne un taux de perte dépendant de l'énergie et de l'angle d'incidence des photons. Ce phénomène se traduit par une aire effective $a_{eff}(k, l, e)$ dépendant de l'énergie e et de la position sur le plan du détecteur (i.e. du pixel k, l). En fait, l'aire effective des télescopes décroît en fonction de la distance au foyer et en fonction de l'énergie pour les photons de plus de 5 keV (cf. figure 1.2).

Parmi les instruments focaux, on compte deux spectromètres à haute résolution RGS (Reflection Grating Spectrometer) pour l'étude des raies d'émission dans les sources ponctuelles et les spectro-imageurs EPIC (European Photon Imaging Camera) pour l'étude des sources étendues,

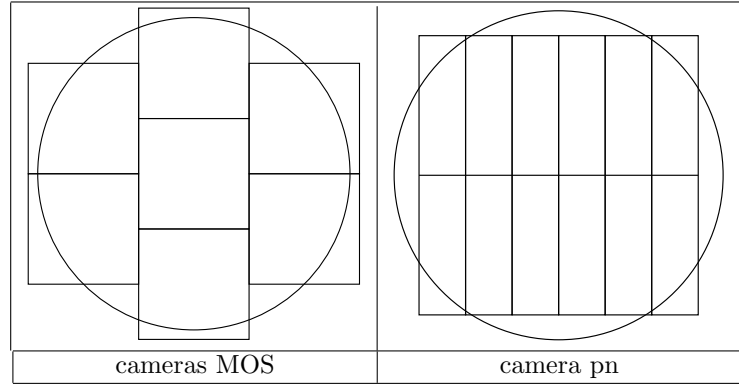


FIG. 1.3 – Géométrie des caméras MOS (7 CCD) et PN (12 CCD). Le cercle matérialise le champ de vue.

avec lesquels nous travaillons. Les trois spectro-imageurs se distinguent en une caméra “pn” et deux caméras MOS (Metal-Oxyde Semi-conductor) de technologies différentes. Ils couvrent un champ de vue de 30 minutes d’arc avec une bande passante de 0.1 à 15 keV. Le chemin optique des rayons lumineux pour les deux télescopes équipés des spectro-imageurs MOS est représenté sur la figure 1.1. Seuls 42 % du flux incident parviennent sur les spectro-imageurs, le reste étant dispersé sur un foyer secondaire où sont placés les spectromètres RGS. Les caméras MOS sont composées d’une mosaïque avec recouvrements de 7 CCD carrés éclairés par l’avant, alors que la caméra pn comprend douze CCD rectangulaires jointifs éclairés par l’arrière (cf. figure 1.3). La caméra pn a un meilleur rendement quantique à haute énergie que les caméras MOS et un temps de lecture bien plus faible, mais aussi une moins bonne résolution spectrale (67 eV à 1 keV contre 57 eV).

Les spectro-imageurs EPIC fonctionnent à chaque pixel comme des compteurs de photons. Chaque photon d’énergie e incident sur le détecteur a une probabilité $QE(e)$ d’être détecté, la fonction $QE(e)$ constituant le rendement quantique de la caméra. A chaque détection est associé un “événement” (k, l, e) caractérisé par l’énergie e du photon incident et la position de son impact sur le plan du détecteur (pixel k, l). En fait, si la majorité des photons ne sont détectés que sur un pixel, de nombreux événements sont constitués de détections sur plusieurs pixels voisins. C’est l’électronique de bord qui calcule la position et l’énergie du photon incident en fonction des déplacement de charges causés par l’impact du photon.

1.2 Nature du signal observé

Le signal observé est une liste de détections de photons ou événements (k, l, e) dont les origines sont multiples. Il y a d’une part les événements réels et d’autre part de fausses détections de photons. Par ailleurs les événements réels peuvent être associés à la source observée ou aux autres sources du champ de vue avec le cas échéant des effets de superposition, – sources ponctuelles, galaxies, .. –.

Dans le cas de l’observation des amas de galaxies les photons issus de la source sont émis par le milieu intra-amas que l’on peut échantillonner en éléments de volume dV . Les n_S photons issus de l’élément de volume dV sont émis avec une énergie e avec une probabilité $S_o(e)$. La fonction normalisée $S_o(e)$ constitue le spectre d’émission du plasma optiquement mince pour l’élément de volume dV . On peut la modéliser en fonction des paramètres du milieu (cf. section 1.2 de la partie I). Le spectre d’émission des photons est ensuite modifié par différents processus au cours du trajet des photons depuis la source jusqu’au détecteur. Le décalage vers le rouge de la source modifie l’énergie e de chaque photon d’un facteur $(\frac{1}{1+\lambda} = \frac{e}{e+hc})$ tandis que l’ionisation de l’hydrogène galactique absorbe la partie des photons dont l’énergie est la plus basse. Par

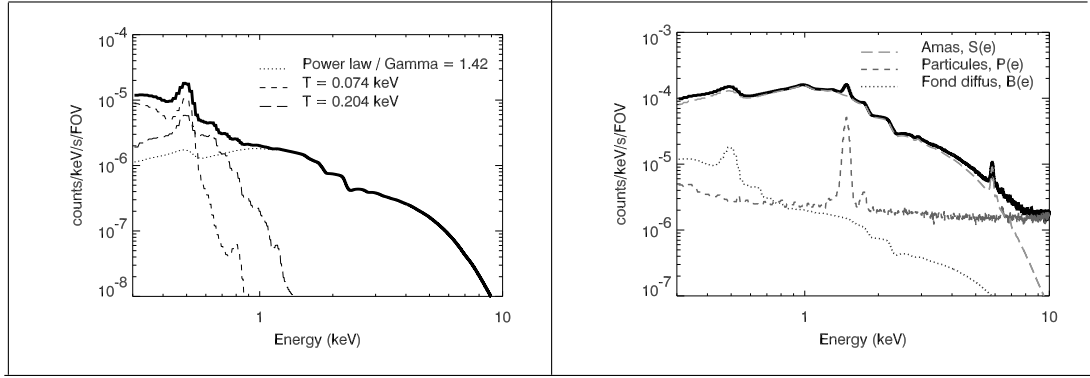


FIG. 1.4 – *Composantes de fond. Gauche: Fond diffus X B(e) à trois composantes (Bulle Locale + CXB). Droite: Emissivité totale F(e) au cours d'une observation d'amas. Fond diffus X B(e), fond particules induites P(e) et amas S(e).*

ailleurs, le spectre observé est la convolution du spectre d'émission théorique $S_o(e)$ par la réponse instrumentale $RI(e)$, dans laquelle intervient le rendement quantique de la caméra ainsi que sa résolution spectrale. Enfin l'imagerie de la source consiste en la projection d'un volume sur le plan du détecteur. Ainsi, puisque le milieu intra-amas est pratiquement optiquement mince, un pixel (k, l) du détecteur reçoit les n_S photons provenant de l'ensemble des éléments de volumes $dV = \delta\Omega(k, l)dr$ contenus dans l'angle solide $\delta\Omega(k, l)$ de la sphère céleste, soit pour l'émissivité d'un amas s'étendant de la distance r_{min} à la distance r_{max} :

$$N_S S(e, k, l) = RI(e, k, l) * \int_{r_{min}}^{r_{max}} n_S(r) S_o(r, e) \delta\Omega(k, l) dr. \quad (1.1)$$

Les autres photons détectés proviennent majoritairement de deux types de sources, des sources ponctuelles (supernovae, galaxies à noyau actif, ..) et un fond diffus émetteur X qui comprend une composante galactique et une composante extragalactique. Les sources ponctuelles les plus brillantes sont reconnaissables par l'allure non thermique de leur spectre d'émission. Une étape du traitement préliminaire des données consiste à sélectionner empiriquement les régions d'émission de ces sources, puis à supprimer les événements associés. Les événements associés au fond diffus ne peuvent par contre pas être isolés spatialement puisqu'il s'agit d'une émission étendue. Il faut donc modéliser cette composante du signal et la prendre en compte dans le cadre d'une analyse multi-composantes. La composante galactique du fond diffus est due à l'émissivité X de restes de supernovae proches (bulle locale, Snowden et al. 1997). Son spectre d'émission $B(e)$ est modélisé par deux composantes thermiques de températures moyennes respectives de 0.204 et 0.074 keV. La composante extragalactique est due au fond diffus X encore appelé CXB (Cosmic X-ray Background), constitué de sources distantes non résolues, essentiellement des galaxies à noyau actif. Le flux de la composante extragalactique est considéré comme uniforme alors que celui des composantes locales dépend de la direction d'observation.

Enfin la principale cause de fausse détection de photons est l'interaction de particules à haute énergie du rayonnement solaire et cosmique avec le détecteur. L'essentiel des particules dues à l'activité solaire sont des protons qui parviennent par bouffées au cours du temps d'exposition. Si cette irrégularité rend difficile une modélisation du spectre des événements associés, elle permet par contre d'isoler les intervalles de temps où l'activité dépasse un certain seuil. Le traitement préliminaire des données comprend une procédure d'exclusion des événements survenus durant ces intervalles de temps. Le reste des particules est essentiellement issu du rayonnement cosmique, dont le flux est cette fois régulier au cours du temps. Le spectre $P(e)$ (cf. figure 1.4) des événements induits par l'interaction du détecteur avec ce rayonnement est pris en compte dans l'analyse.

Dans le cas de l'observation des amas et groupes de galaxies, le signal analysé est finalement, après soustraction des événements associés aux sursauts d'activité solaire et aux sources ponctuelles, une liste de photons (k,l,e) provenant respectivement:

- de l'amas source;
- du fond diffus X, avec la bulle locale et le CXB;
- du fond de particules induit par les rayons cosmiques.

Pour exprimer la probabilité $F(k,l,e)$ d'observation d'un photon d'énergie e dans le pixel k,l , il faut connaître pour chaque composante du signal, la probabilité d'émission d'un photon, $f_i(k,l,e)$, mais aussi le nombre de photons émis n_i . Dans notre cas le modèle d'émission $F(k,l,e)$ comprend la composante thermique d'émissivité de l'amas, de spectre d'émission $S(e)$ et les composantes de fond diffus et particules, de spectre d'émission $B(e)$ et $P(e)$ respectivement. La probabilité $F(k,l,e)$ s'exprime en fonction des spectres ou probabilités d'émission $S(e), P(e), B(e)$ et des émissivités respectives N_S, N_P et N_B de chaque composante(cf. figure 1.4):

$$N_F F(k,l,e) = N_S S(e) + N_P P(e) + N_B B(e). \quad (1.2)$$

Chapitre 2

La transformée en ondelettes.

2.1 L'analyse temps-fréquence et la transformée en ondelettes.

Les algorithmes d'imagerie et de spectro-imagerie décrits dans cette thèse ont pour point commun de chercher à cartographier la valeur d'un paramètre du milieu intra-amas tel qu'observé sur le plan du ciel. Pour cela, ils cherchent à restituer au mieux les variations locales plus ou moins étendues de ce paramètre à partir d'un signal bruité en dépendant. L'étude des variations spatiales du signal fait intervenir des filtres passe-haut. L'analyse simultanée de la localisation et de l'extension de ces variations fait appel à l'analyse temps-fréquence pour les signaux temporels, à l'analyse spatio-fréquentielle pour l'imagerie. C'est dans ce contexte que l'utilisation de la transformation en ondelettes s'insère et devient naturelle.

Une ondelette ψ est une fonction normalisée, $\|\psi\| = 1$ et de moyenne nulle, $\int_{-\infty}^{+\infty} \psi(t) dt = 0$, caractérisée par un paramètre d'échelle s et un paramètre de position u (Grossman & Morlet, 1984):

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right). \quad (2.1)$$

La transformée en ondelettes continue $W_f(u,s)$ d'un signal temporel $f(t)$ à une dimension exprime la corrélation de f avec l'ondelette ψ . Il s'agit d'une transformation linéaire et invariante par translation, où l'ondelette se comporte comme un filtre passe-haut d'analyse des variations du signal f aux différentes échelles s et positions u de l'espace temps-fréquence:

$$W_f(u,s) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-u}{s}\right) dt. \quad (2.2)$$

Ainsi définie, la transformée en ondelettes est une projection du signal $f(t) \in \mathbb{R}$ à une dimension dans l'espace temps-fréquence $\mathbf{L}^2(\mathbb{R})$ à deux dimensions. Il s'agit donc d'une transformée redondante. D'autant que le principe d'incertitude stipule que l'énergie d'une fonction et de sa transformée de Fourier ne peuvent être simultanément concentrées sur des intervalles arbitrairement petits. En d'autre termes on ne peut localiser l'énergie d'un signal en tout point de l'espace temps-fréquence, mais seulement dans des "rectangles d'Heisenberg" d'aire non arbitrairement petite. Pour supprimer la redondance de la transformée, il faut donc projeter le signal $f(t)$ sur un espace temps-fréquence échantillonné en rectangles d'Heisenberg. Ce type de projection peut être effectué avec des ondelettes spécifiques, dont les paramètres de temps et de fréquence sont discrétisés et qui ont la propriété de constituer des bases orthonormales de l'espace temps-fréquence. L'échantillonnage dyadique de l'axe des échelles, c'est-à-dire une discrétisation où les échelles de la transformée sont des multiples de 2 successifs, $\{s = 2^j\}_{j \in \mathbb{Z}}$, permet par exemple de construire les ondelettes orthogonales $\psi_{j,n}$, telles que:

$$\left\{ \psi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t-2^j n}{2^j}\right) \right\}_{(j,n) \in \mathbb{Z}^2}. \quad (2.3)$$

Et le signal f s'exprime alors par une somme discrète de vecteurs de bases de l'espace temps-fréquence $\psi_{j,n}$ pondérées par des coefficients en ondelettes $\langle f, \psi_{j,n} \rangle$:

$$f = \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n}. \quad (2.4)$$

2.2 Analyse multirésolution

Une analyse en ondelettes sur une base orthogonale de type (2.3) constitue la réunion d'un ensemble de "signaux de détails" D_{2^j} caractérisant les variations locales du signal f à toutes les résolutions, $\{s = 2^j\}_{j \in \mathbb{Z}}$:

$$D_{2^j-1} = \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n}. \quad (2.5)$$

$$f = \sum_{j=-\infty}^{+\infty} D_{2^j}. \quad (2.6)$$

Cette analyse peut être tronquée en fixant une résolution maximale $\{s = 2^L\}$, on obtient alors une approximation de f à la résolution 2^L :

$$A_{2^j} = \sum_{j=L}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n}. \quad (2.7)$$

En combinant (2.6) et (2.7) on peut écrire:

$$\begin{aligned} A_{2^j} &= \sum_{j=L}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n} \\ &= \sum_{j=L}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle f, \psi_{j,n} \rangle \psi_{j,n} + \langle f, \psi_{L,n} \rangle \psi_{L,n} \\ &= A_{2^{L-1}} + D_{2^{L-1}} \end{aligned} \quad (2.8)$$

On en déduit d'une part, que l'approximation A_{2^j} de f contient toute l'information contenue dans l'approximation $A_{2^{j-1}}$ à plus basse résolution et que d'autre part, la perte d'information subie en approchant A_{2^j} par $A_{2^{j-1}}$ est contenue dans le signal détail $D_{2^{j-1}} = A_{2^j} - A_{2^{j-1}}$. Par récurrence, on en déduit que la transformée en ondelettes de f (2.6)) constitue l'ensemble des informations perdues en approchant f par des lissages aux résolutions successives $\{s = 2^j\}_{j \in \mathbb{Z}}$. Il existe ainsi une dualité entre les signaux détails D_{2^j} de la transformée en ondelettes du signal f discrétisée sur une base orthogonale de l'espace temps-fréquence $\mathbf{L}^2(\mathbb{R})$ et la suite des approximations multirésolution A_{2^j} de ce signal.

2.2.1 Sous-espaces d'approximation multirésolution

L'analyse multirésolution définit des sous-espaces emboîtés \mathbf{V}_j d'approximations de $\mathbf{L}^2(\mathbb{R})$ aux résolutions successives. La dualité entre approximations et signaux détails se formalise dans ce cadre en définissant un sous-espace de détails \mathbf{U}_j , dual de chaque sous-espace d'approximation \mathbf{V}_j .

Les propriétés des sous-espaces d'approximation, \mathbf{V}_j , sont les suivantes:

- Invariance par translation,

$$\forall (j, k) \in \mathbb{Z}^2, f(t) \in \mathbf{V}_j \Leftrightarrow f(t - 2^j k) \in \mathbf{V}_j; \quad (2.9)$$

- Causalité des informations entre sous-espaces de résolutions successives,

$$\forall j \in \mathbb{Z} \mathbf{V}_{j+1} \subset \mathbf{V}_j; \quad (2.10)$$

- Invariance par dilatation entre sous-espaces de résolutions successives,

$$\forall j \in \mathbb{Z}, f(t) \in \mathbf{V}_j \Leftrightarrow f\left(\frac{t}{2}\right) \in \mathbf{V}_{j+1}; \quad (2.11)$$

- Perte de toute information à la résolution nulle,

$$\lim_{j \rightarrow +\infty} \mathbf{V}_j = \bigcap_{j=-\infty}^{+\infty} \mathbf{V}_j = \{0\}; \quad (2.12)$$

- Complétude de l'approximation et convergence de l'approximation vers le signal à la résolution infinie,

$$\lim_{j \rightarrow -\infty} \mathbf{V}_j = \bigcup_{j=-\infty}^{+\infty} \mathbf{V}_j = \mathbf{L}^2(\mathbb{R}); \quad (2.13)$$

- Existence d'une base discrète de \mathbf{V}_0 , $\{\beta(t-n)\}_{n \in \mathbb{Z}}$.

Chaque sous-espace de détails W_j est quant-à-lui défini comme le complément orthogonal du sous-espace d'approximation V_j , dans le sous-espace d'approximation à la résolution supérieure, V_{j-1} :

$$\mathbf{V}_{j-1} = \mathbf{V}_j \oplus \mathbf{U}_j; \quad (2.14)$$

Et l'espace temps-fréquence $\mathbf{L}^2(\mathbb{R})$ est finalement engendré par la réunion des bases des sous-espaces de détails W_j à toutes les résolutions:

$$\mathbf{L}^2(\mathbb{R}) = \bigoplus_{j=-\infty}^{+\infty} \mathbf{U}_j; \quad (2.15)$$

2.2.2 Fonction d'échelle

Les sous-espaces duaux \mathbf{V}_j et \mathbf{U}_j ont été construits pour être associés à des bases orthogonales. La base du sous-espace de détail \mathbf{U}_j est une base d'ondelettes orthogonales $\{\psi_{j,n}\}_{j,n \in \mathbb{Z}^2}$ du type (2.3). On montre que la base du sous-espace d'approximation \mathbf{V}_j est une base de *fonctions d'échelles* $\{\phi_{j,n}\}_{n \in \mathbb{Z}}$ du type:

$$\left\{ \phi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \phi\left(\frac{t-n}{2^j}\right) \right\}_{(n) \in \mathbb{Z}}. \quad (2.16)$$

L'approximation de f à la résolution j s'exprime ainsi par:

$$A_j = \sum_{n=-\infty}^{+\infty} \langle f, \phi_{j,n} \rangle \phi_{j,n}. \quad (2.17)$$

Où les produits scalaires $\langle f, \phi_{j,n} \rangle$ expriment la corrélation de f avec la fonction d'échelle, $\{\phi_{j,n}\}_{j,n \in \mathbb{Z}^2}$, de même que les coefficients en ondelettes $\langle f, \psi_{j,n} \rangle$ expriment la corrélation de f avec l'ondelette $\{\psi_{j,n}\}_{j,n \in \mathbb{Z}^2}$ (cf. equation 2.2):

$$\langle f, \phi_{j,n} \rangle = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{2^j}} \phi\left(\frac{t-2^j n}{2^j}\right) dt. \quad (2.18)$$

La fonction d'échelle $\{\phi_{j,n}\}_{n \in \mathbb{Z}}$ se comporte donc à chaque résolution comme un filtre passe-bas d'analyse de f . Ce filtre constitue le filtre dual de l'ondelette $\{\psi_{j,n}\}_{n \in \mathbb{Z}}$, filtre passe-haut d'analyse de f .

2.2.3 Relations inter-échelles et filtrage

Le formalisme de l'analyse multirésolution permet d'exprimer des relations entre les bases d'approximations de résolutions successives, \mathbf{V}_{j+1} et \mathbf{V}_j . Ainsi la propriété de causalité (2.10) des sous-espaces d'approximation implique que $\mathbf{V}_1 \subset \mathbf{V}_0$ et permet d'exprimer les vecteurs de base de \mathbf{V}_1 , $\frac{1}{\sqrt{2}}\phi(\frac{t}{2})$, dans la base de \mathbf{V}_0 , $\{\phi(t-n)_{j,n \in \mathbb{Z}^2}\}$. Cette projection, aussi appelée équation de dilatation, exprime la dilatation de ϕ par 2 en fonction de ses translations entières à la résolution supérieure:

$$\begin{aligned} \frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right) &= \sum_{n=-\infty}^{+\infty} \left\langle \frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right), \phi(t-n) \right\rangle \phi(t-n), \\ &= \sum_{n=-\infty}^{+\infty} h[n]\phi(t-n). \end{aligned} \quad (2.19)$$

De même, les vecteurs de base du sous-espace de détail \mathbf{U}_1 s'expriment dans la base de \mathbf{V}_1 :

$$\begin{aligned} \frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right) &= \sum_{n=-\infty}^{+\infty} \left\langle \frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right), \psi(t-n) \right\rangle \phi(t-n), \\ &= \sum_{n=-\infty}^{+\infty} g[n]\phi(t-n). \end{aligned} \quad (2.20)$$

Les projecteurs orthogonaux $h[n] = \left\langle \frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right), \phi(t-n) \right\rangle$ et $g[n] = \left\langle \frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right), \psi(t-n) \right\rangle$, de \mathbf{V}_1 et \mathbf{U}_1 dans \mathbf{V}_0 définissent des relations de filtrage linéaire entre l'approximation \mathbf{V}_0 d'une part et les sous-espaces d'approximation et de détails, \mathbf{V}_1 et \mathbf{U}_1 , d'autre part. Ce sont ces relations de filtrage linéaire inter-échelles qui sont utilisées pour construire des algorithmes rapides de calcul des coefficients en ondelettes dans une base d'ondelettes orthogonale $\{\psi_{j,n}\}_{j,n \in \mathbb{Z}^2}$ du type (2.3). La relation (2.20) exprime le couplage entre fonctions d'échelles et ondelettes en fonction du filtre g , tandis que la dépendance entre les filtres h et g est déduite de l'expression de \mathbf{U}_j comme complément orthogonal de \mathbf{V}_j . Ces contraintes s'expriment dans l'espace de Fourier par les conditions suivantes (Mallat, 1989):

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2}}\hat{g}\left(\frac{\omega}{2}\right)\hat{\phi}\left(\frac{\omega}{2}\right) \quad (2.21)$$

$$\hat{g}(\omega) = e^{-i\omega}\hat{h}^*(\omega + \pi). \quad (2.22)$$

Enfin la transformée de Fourier inverse de (2.22) donne la relation suivante entre filtres discrets h et g :

$$g[n] = (-1)^{1-n}h[1-n] \quad (2.23)$$

2.2.4 Analyse multirésolution et transformée en ondelettes à deux dimensions

Nous avons jusqu'ici analysé un signal $f(t)$ à une dimension en le projetant dans l'espace temps-fréquence $\mathbf{L}^2(\mathbb{R})$ à deux dimensions. La transformée en ondelettes à deux dimensions permet de projeter un signal à deux dimensions, par exemple une image $I(k,l)$, dans l'espace spatio-fréquentiel $\mathbf{L}^2(\mathbb{R}^2)$ à quatre dimensions. Pour cela on peut étendre le formalisme de l'analyse multirésolution, en définissant pour chaque résolution j des sous-espaces séparables \mathbf{V}_j^2 à 2D, produits tensoriels de sous-espaces multirésolution à 1D:

$$\mathbf{V}_j^2 = \mathbf{V}_j \otimes \mathbf{V}_j. \quad (2.24)$$

Les sous-espaces de détails \mathbf{U}_j^2 sont les compléments orthogonaux des sous-espaces d'approximation \mathbf{V}_j^2 dans \mathbf{V}_{j-1}^2 :

$$\mathbf{V}_{j-1}^2 = \mathbf{V}_j^2 \oplus \mathbf{U}_j^2 \quad (2.25)$$

En utilisant la complémentarité des sous-espaces à 1D, $\mathbf{V}_{j-1} = \mathbf{V}_j \oplus \mathbf{U}_j$, on peut alors écrire:

$$\begin{aligned} \mathbf{V}_j^2 \oplus \mathbf{U}_j^2 &= \mathbf{V}_{j-1} \otimes \mathbf{V}_{j-1} \\ \mathbf{V}_j^2 \oplus \mathbf{U}_j^2 &= (\mathbf{V}_j \oplus \mathbf{U}_j) \otimes (\mathbf{V}_j \oplus \mathbf{U}_j) \\ \mathbf{V}_j^2 \oplus \mathbf{U}_j^2 &= \mathbf{V}_j^2 \oplus (\mathbf{V}_j \otimes \mathbf{U}_j) \oplus (\mathbf{U}_j \otimes \mathbf{V}_j) \oplus (\mathbf{U}_j \otimes \mathbf{U}_j) \\ \mathbf{U}_j^2 &= (\mathbf{V}_j \otimes \mathbf{U}_j) \oplus (\mathbf{U}_j \otimes \mathbf{V}_j) \oplus (\mathbf{U}_j \otimes \mathbf{U}_j) \end{aligned} \quad (2.26)$$

Les sous-espaces de détails \mathbf{U}_j^2 sont donc constitués de la somme directe des trois sous-espaces $(\mathbf{V}_j \otimes \mathbf{U}_j)$, $(\mathbf{U}_j \otimes \mathbf{V}_j)$ et $(\mathbf{U}_j \otimes \mathbf{U}_j)$. On en déduit que leurs bases orthonormées sont engendrées par la réunion de trois bases d'ondelettes, $\{\psi_{j,k,l}^1\}_{k,l \in \mathbb{Z}^2}$, $\{\psi_{j,k,l}^2\}_{k,l \in \mathbb{Z}^2}$, $\{\psi_{j,k,l}^3\}_{k,l \in \mathbb{Z}^2}$, construites à partir des bases orthonormées $\{\phi_{j,n}\}_{n \in \mathbb{Z}}$ et $\{\psi_{j,n}\}_{n \in \mathbb{Z}}$, de \mathbf{V}_j et \mathbf{U}_j :

$$\begin{aligned} \{\psi_{j,k,l}^1(k,l)\}_{k,l \in \mathbb{Z}^2} &= \{\phi_{j,k}(k)\psi_{j,l}(l)\}_{k,l \in \mathbb{Z}^2} \\ \{\psi_{j,k,l}^2(k,l)\}_{k,l \in \mathbb{Z}^2} &= \{\psi_{j,k}(k)\phi_{j,l}(l)\}_{k,l \in \mathbb{Z}^2} \\ \{\psi_{j,k,l}^3(k,l)\}_{k,l \in \mathbb{Z}^2} &= \{\psi_{j,k}(k)\psi_{j,l}(l)\}_{k,l \in \mathbb{Z}^2}. \end{aligned} \quad (2.27)$$

Les ondelettes $\psi_{j,k,l}^1(k,l)$ et $\psi_{j,k,l}^2(k,l)$ associées aux sous-espaces $(\mathbf{V}_j \otimes \mathbf{U}_j)$ et $(\mathbf{U}_j \otimes \mathbf{V}_j)$ se comportent comme des filtres passe-haut orientés respectivement sur les axes horizontal et vertical de l'image. L'information contenue dans les sous-espaces $(\mathbf{V}_j \otimes \mathbf{U}_j)$ et $(\mathbf{U}_j \otimes \mathbf{V}_j)$ est donc essentiellement et respectivement constituée des structures horizontales et verticales de l'image. Enfin l'ondelette $\psi_{j,k,l}^3(k,l)$ associée au sous-espace $(\mathbf{U}_j \otimes \mathbf{U}_j)$ peut être isotrope ou anisotrope suivant la parité de l'ondelette à 1D $\psi_{j,n}(n)$. Dans le cas anisotrope, les détails analysés dans le sous-espace $(\mathbf{U}_j \otimes \mathbf{U}_j)$ sont essentiellement constitués des structures diagonales de l'image. L'ensemble des bases orthogonales des sous-espaces \mathbf{U}_j^2 permettent finalement de définir une transformée en ondelettes à deux dimensions en engendrant l'espace $\mathbf{L}^2(\mathbb{R}^2)$:

$$\mathbf{L}^2(\mathbb{R}^2) = \oplus_{j=-\infty}^{+\infty} \mathbf{U}_j^2. \quad (2.28)$$

Les relations inter-échelles entre bases se déduisent de (2.19) et (2.20) en utilisant la séparabilité des ondelettes sur les deux axes, ainsi on peut exprimer l'équation de dilatation des fonctions d'échelles à deux dimensions:

$$\begin{aligned} \frac{1}{2}\phi\left(\frac{k}{2}\right)\phi\left(\frac{l}{2}\right) &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \left\langle \frac{1}{2}\phi\left(\frac{k}{2}\right)\phi\left(\frac{l}{2}\right), \phi(t-m)\phi(t-n) \right\rangle \phi(t-m)\phi(t-n), \\ &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} h[m]h[n]\phi(t-m)\phi(t-n). \end{aligned} \quad (2.29)$$

Et de la même façon, on exprime les bases des trois sous-espaces de détails en fonction de la base du sous-espace d'approximation à la résolution supérieure:

$$\begin{aligned}
\frac{1}{2}\psi\left(\frac{k}{2}\right)\phi\left(\frac{l}{2}\right) &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} g[m]h[n]\phi(t-m)\phi(t-n) \\
\frac{1}{2}\phi\left(\frac{k}{2}\right)\psi\left(\frac{l}{2}\right) &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} h[m]g[n]\phi(t-m)\phi(t-n) \\
\frac{1}{2}\psi\left(\frac{k}{2}\right)\psi\left(\frac{l}{2}\right) &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} g[m]g[n]\phi(t-m)\phi(t-n)
\end{aligned} \tag{2.30}$$

Il existe d'autres types de bases orthogonales permettant de construire des transformées en ondelettes à deux dimensions. L'intérêt de celle-ci est d'être engendrée par des bases séparables sur les deux axes (k, l) , ce qui permet de définir des algorithmes de filtrage séparables en filtrages à 1D, plus rapides que les algorithmes à 2D.

2.3 Algorithmes de calcul des transformées en ondelettes discrètes

2.3.1 Algorithme multirésolution

L'algorithme multirésolution est un algorithme de calcul de transformée en ondelettes orthogonales d'un signal échantillonné. Cet algorithme repose sur l'analyse du signal échantillonné $f(t)$, considéré comme une approximation de résolution finie L d'un signal avec des détails aussi fins que possibles. On en détaille ici d'une part, le processus de *décomposition* du signal échantillonné $f(t)$ en sa transformée en ondelettes $W(t, j)$, d'autre part le processus inverse de *reconstruction* du signal à partir de sa transformée.

Processus de décomposition du signal

Etant donné un signal $f(t)$ échantillonné à la résolution L , on calcule dans un premier temps les coefficients d'approximations $F_j[n]$ du signal à toutes les résolutions $J < j \leq L$, entre une résolution minimale J et la résolution du signal L . Pour cela, on procède à une succession de filtrages passe-bas des approximations. Dans un deuxième temps on procède au filtrage passe-haut des coefficients d'approximations et on en déduit les coefficients de détails $W_j[n]$ ou coefficients en ondelettes.

Les coefficients d'approximations $F_j[n]$ et de détails $W_j[n]$ s'expriment comme les projecteurs de $f(t)$ sur \mathbf{V}_j et \mathbf{U}_j respectivement ¹:

$$\begin{aligned}
F_j[n] &= \langle f, \phi_{j,n} \rangle \\
W_j[n] &= \langle f, \psi_{j,n} \rangle.
\end{aligned} \tag{2.31}$$

Pour déterminer les relations de filtrage, on décompose les bases des sous-espaces \mathbf{V}_{j+1} et \mathbf{U}_{j+1} dans la base de \mathbf{V}_j :

$$\begin{aligned}
\phi_{j+1,p} &= \sum_{n=-\infty}^{+\infty} \langle \phi_{j+1,p}, \phi_{j,n} \rangle \phi_{j,n} \\
\psi_{j+1,p} &= \sum_{n=-\infty}^{+\infty} \langle \psi_{j+1,p}, \phi_{j,n} \rangle \phi_{j,n}.
\end{aligned} \tag{2.32}$$

1. On appellera en principe coefficients ou signaux d'approximations et de détails, $F_j[n]$ et $W_j[n]$, les projecteurs de f sur les sous-espaces d'approximations et de détails. Par abus de langage, on les confondra parfois avec les approximations A_j et signaux détails D_j qui en sont les projections.

Le produit scalaire $\langle \phi_{j+1,p}, \phi_{j,n} \rangle$ se réécrit en fonctions du filtre h en utilisant la relation inter-échelles (2.19) à la résolution $j = 0$. On montre alors par changement de variables qu'il est invariant quelle que soit la résolution:

$$\begin{aligned} \langle \phi_{j+1,p}, \phi_{j,n} \rangle &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2^{j+1}}} \phi\left(\frac{t - 2^{j+1}p}{2^{j+1}}\right) \frac{1}{\sqrt{2^j}} \phi^*\left(\frac{t - 2^j n}{2^j}\right) dt \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2}} \phi\left(\frac{t}{2}\right) \phi^*(t - n + 2p) dt \\ &= \langle \frac{1}{\sqrt{2}} \phi\left(\frac{t}{2}\right) \phi(t - n + 2p) \rangle \\ &= h[n - 2p] \end{aligned} \tag{2.33}$$

$$\tag{2.34}$$

De même pour le produit scalaire $\langle \psi_{j+1,p}, \phi_{j,n} \rangle$ avec (2.20):

$$\begin{aligned} \langle \psi_{j+1,p}, \phi_{j,n} \rangle &= \langle \frac{1}{\sqrt{2}} \psi\left(\frac{t}{2}\right) \phi(t - n + 2p) \rangle \\ &= g[n - 2p]. \end{aligned} \tag{2.35}$$

En projetant f sur les sous-espaces d'approximation, on déduit de (2.32) les relations suivantes:

$$\begin{aligned} \langle f, \phi_{j+1,p} \rangle &= \sum_{n=-\infty}^{+\infty} h[n - 2p] \langle f, \phi_{j,n} \rangle \\ \langle f, \psi_{j+1,p} \rangle &= \sum_{n=-\infty}^{+\infty} g[n - 2p] \langle f, \phi_{j,n} \rangle, \end{aligned} \tag{2.36}$$

que l'on peut réécrire en relations entre coefficients d'approximations $F_j[n]$ d'une part, de détails $W_{j-1}[n]$ et d'approximations $F_{j-1}[n]$ à la résolution inférieure d'autre part:

$$\begin{aligned} F_{j+1}[p] &= \sum_{n=-\infty}^{+\infty} h[n - 2p] F_j[n], \\ W_{j+1}[p] &= \sum_{n=-\infty}^{+\infty} g[n - 2p] F_j[n] \end{aligned} \tag{2.37}$$

Ce filtrage revient à convoluer les coefficients d'approximation $F_j[n]$ par les filtre $h[-2p] = \bar{h}[2p]$ et $g[-2p] = \bar{g}[2p]$. Les signaux sont donc convolués avec un pas double de leur pas d'échantillonnage, et le pas d'échantillonnage de $F_{j+1}[n]$ est deux fois plus grand que le pas d'échantillonnage de $F_j[n]$. En conséquence, si le signal d'approximations $F_j[n]$ est borné et contient un nombre fini d'échantillons, N , le signal $F_{j+1}[n]$ ne contient que $\frac{N}{2}$ échantillons. L'algorithme multirésolution implique donc une *décimation* dyadique du nombre d'échantillons à chaque échelle. Finalement, le processus de *décomposition* du signal $f(t)$ en sa transformée en ondelettes se ramène à une succession de filtrages complémentaires passe-haut et passe-bas de ses approximations successives F_j où le nombre d'échantillons est décimé d'un facteur 2 à chaque filtrage (cf. figure 2.1). Notons que le processus de décimation implique que le nombre total d'échantillons N est conservé d'échelle en échelle entre, d'une part, le signal d'approximations $F_j[n]$ et d'autre part, la somme des signaux complémentaires de détails W_{j+1} et d'approximations F_{j+1} à la résolution inférieure.

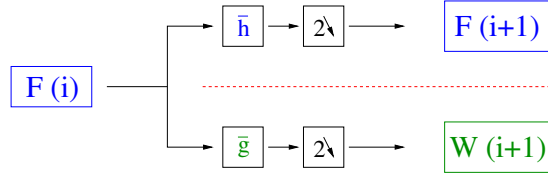


FIG. 2.1 – Décomposition de du signal d'approximations $F(i)$ en un signal d'approximation $F(i+1)$ et un signal de détails $W(i+1)$. Les deux signaux complémentaires $F(i)$ et $W(i+1)$ sont obtenus par filtrages passe-haut et passe-bas de $F(i)$ et décimation d'un facteur 2 du nombre d'échantillons.

Processus de reconstruction de la transformée

Examinons maintenant le processus inverse de reconstruction du signal $f(t)$ à partir de sa transformée en ondelettes. Les coefficients d'approximation successifs F_j sont maintenant calculés de la plus basse résolution J à celle du signal L , à partir des signaux de détails W_{j+1} et d'approximations F_{j+1} aux résolutions précédentes. Les relations de filtrage inter-échelles se déduisent de la décomposition de la base de \mathbf{V}_j dans la réunion des bases des compléments orthogonaux, \mathbf{V}_{j+1} et \mathbf{U}_{j+1} . Elles font intervenir les produits scalaires $\langle \phi_{j,p}, \phi_{j+1,n} \rangle$ et $\langle \phi_{j,p}, \psi_{j+1,n} \rangle$ que l'on peut réécrire en fonction des filtres h et g en utilisant (2.34) et (2.35):

$$\begin{aligned} \phi_{j,p} &= \sum_{n=-\infty}^{+\infty} \langle \phi_{j,p}, \phi_{j+1,n} \rangle \phi_{j+1,n} + \sum_{n=-\infty}^{+\infty} \langle \phi_{j,p}, \psi_{j+1,n} \rangle \psi_{j+1,n} \\ &= \sum_{n=-\infty}^{+\infty} h[p-2n] \phi_{j+1,n} + \sum_{n=-\infty}^{+\infty} g[p-2n] \psi_{j+1,n}. \end{aligned} \quad (2.38)$$

La formule de *reconstruction* à la résolution j s'exprime alors par:

$$F_j[p] = \sum_{n=-\infty}^{+\infty} h[p-2n] F_{j+1}[n] + \sum_{n=-\infty}^{+\infty} g[p-2n] W_{j+1}[n]. \quad (2.39)$$

Le signal d'approximations $F_j[p]$ est donc obtenu en additionnant les signaux d'approximation $F_{j+1}[p]$ et de détails $W_{j+1}[p]$, convolués par les filtres $h[p/2]$ et $g[p/2]$. Les signaux F_{j+1} et W_{j+1} sont donc convolués avec un demi-pas d'échantillonnage. En pratique, la convolution doit être effectuée sur des signaux suréchantillonnés d'un facteur 2. Une interpolation simple consiste à intercaler un zéro entre chaque échantillon des signaux $h[p/2]$ et $g[p/2]$ (cf. figure 2.2).

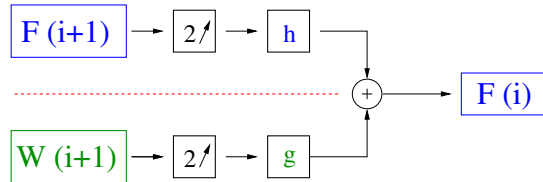


FIG. 2.2 – Reconstruction du signal d'approximation $F(i)$ à partir du signal d'approximation $F(i+1)$ et du signal de détails $W(i+1)$. Les signaux $F(i+1)$ et $W(i+1)$ sont suréchantillonnés d'un facteur 2 en intercalant des zéros entre chaque échantillon, convolués par les filtres de reconstruction h et g , puis additionnés.

Algorithme multirésolution à deux dimensions

L'algorithme multirésolution à deux dimensions repose sur l'analyse multirésolution en bases séparables à deux dimensions (cf. 2.2.4). A chaque résolution, on calcule maintenant un signal d'approximation F_j et ses trois signaux de détails complémentaires, $W_{h,j}$, $W_{v,j}$ et $W_{d,j}$, associés respectivement aux structures horizontales, verticales et diagonales. Ces signaux sont déterminés par projection du signal $I(k,l)$ à la résolution L sur les sous-espaces \mathbf{V}_j^2 , $(\mathbf{V}_{j+1} \otimes \mathbf{U}_{j+1})$, $(\mathbf{U}_{j+1} \otimes \mathbf{V}_{j+1})$ et $(\mathbf{U}_{j+1} \otimes \mathbf{U}_{j+1})$:

$$\begin{aligned} F_j[k,l] &= \langle I(k,l), \phi(k)\phi(l) \rangle \\ W_{h,j}[k,l] &= \langle I(k,l), \phi(k)\psi(l) \rangle \\ W_{v,j}[k,l] &= \langle I(k,l), \psi(k)\phi(l) \rangle \\ W_{d,j}[k,l] &= \langle I(k,l), \psi(k)\psi(l) \rangle. \end{aligned} \quad (2.40)$$

Comme dans l'algorithme à une dimension, les signaux de détails et d'approximation sont calculés par filtrages du signal d'approximation à la résolution supérieure. Les relations de décomposition se déduisent des relations inter-échelles entre bases (2.29) à (2.30):

$$\begin{aligned} F_{j+1}[k,l] &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} h[m-2k]h[n-2l]F_j[m,n], \\ W_{h,j+1}[k,l] &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} h[m-2k]g[n-2l]F_j[m,n], \\ W_{v,j+1}[k,l] &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} g[m-2k]h[n-2l]F_j[m,n], \\ W_{d,j+1}[k,l] &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} g[m-2k]g[n-2l]F_j[m,n]. \end{aligned} \quad (2.41)$$

Au cours de la décomposition du signal $I(k,l)$, les signaux successifs d'approximations du signal $F_j[m,n]$ sont donc convolués par les filtres $h[2k]h[2l]$, $h[2k]g[2l]$, $g[2k]h[2l]$ et $g[2k]g[2l]$. Le pas de convolution étant deux fois plus large que le pas d'échantillonnage, on procède à une décimation d'un facteur 2 sur chaque axe des signaux de détail et d'approximation (cf. figure 2.3). Si le signal d'approximation $F_{j+1}[k,l]$ contient N échantillons, les signaux de détails et approximations, F_{j+1} , $W_{h,j+1}$, $W_{v,j+1}$ et $W_{d,j+1}$ contiennent chacun $\frac{N}{4}$ échantillons. Le nombre total d'échantillons est conservé d'échelle en échelle comme dans l'algorithme à une dimension.

La relations de reconstruction se déduisent en exprimant la base de \mathbf{V}_j dans les bases des sous-espaces $(\mathbf{V}_{j+1} \otimes \mathbf{U}_{j+1})$, $(\mathbf{U}_{j+1} \otimes \mathbf{V}_{j+1})$ et $(\mathbf{U}_{j+1} \otimes \mathbf{U}_{j+1})$, puis en y projetant le signal $I(k,l)$:

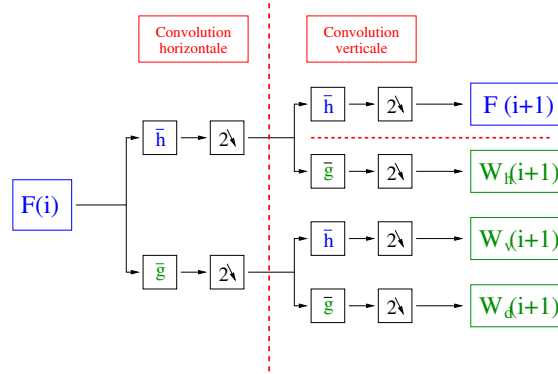


FIG. 2.3 – Décomposition de l'approximation $F(i)$ en un signal d'approximation $F(i+1)$ et trois signaux de détails $W_h(i+1)$, $W_v(i+1)$ et $W_d(i+1)$. Les quatre signaux à la résolution $i+1$ sont obtenus par une combinaison de filtrages et de décimations d'un facteur 2 sur les deux axes.

$$\begin{aligned}
 \phi_{j,k}\phi_{j,l} &= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle \phi_{j,k}\phi_{j,l}, \phi_{j+1,m}\phi_{j+1,n} \rangle \phi_{j+1,m}\phi_{j+1,n} \\
 &+ \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle \phi_{j,k}\phi_{j,l}, \psi_{j+1,m}\phi_{j+1,n} \rangle \psi_{j+1,m}\phi_{j+1,n} \\
 &+ \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle \phi_{j,k}\phi_{j,l}, \phi_{j+1,m}\psi_{j+1,n} \rangle \phi_{j+1,m}\psi_{j+1,n} \\
 &+ \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle \phi_{j,k}\phi_{j,l}, \psi_{j+1,m}\psi_{j+1,n} \rangle \psi_{j+1,m}\psi_{j+1,n}, \quad (2.42)
 \end{aligned}$$

$$\begin{aligned}
 F_j[k,l] &= \sum_{n=-\infty}^{+\infty} h[k-2m]h[l-2n]F_{j+1}[m,n] \\
 &+ \sum_{n=-\infty}^{+\infty} h[k-2m]g[l-2n]W_{h,j+1}[m,n] \\
 &+ \sum_{n=-\infty}^{+\infty} g[k-2m]h[l-2n]W_{v,j+1}[m,n] \\
 &+ \sum_{n=-\infty}^{+\infty} g[k-2m]g[l-2n]W_{d,j+1}[m,n]. \quad (2.43)
 \end{aligned}$$

Les signaux d'approximation sont donc reconstruits à partir de la somme des signaux de détails et d'approximation à la résolution inférieure, convolués successivement sur les deux axes par des combinaisons de filtres $h[p/2]$ et $g[p/2]$. Comme pour l'algorithme à une dimension, le pas de convolution sur chaque axe est deux fois plus petit que le pas d'échantillonnage et il faut suréchantillonner les signaux en insérant des zéros (cf. figure 2.4).

Algorithme multirésolution appliqué à un signal borné

L'analyse multirésolution à une ou à deux dimensions permet donc de construire un algorithme rapide de calcul d'une transformée en ondelettes orthogonale. L'algorithme reste rapide à deux dimension grâce à l'utilisation de bases séparables sur les deux axes. La transformée d'un signal

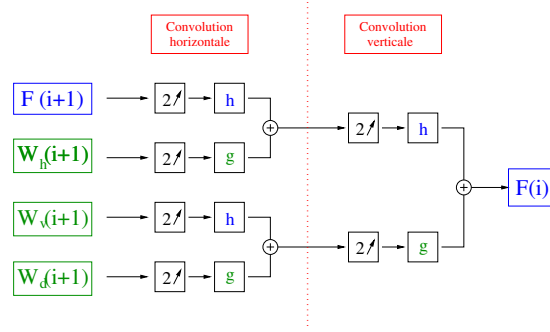


FIG. 2.4 – Reconstruction de l'approximation $F(i)$ à partir des signaux de détails $W_h(i+1)$, $W_v(i+1)$ et $W_d(i+1)$ et d'approximation $F(i+1)$. Les signaux $W_h(i+1)$, $W_v(i+1)$ et $W_d(i+1)$ et $F(i+1)$ sont convolués séparément sur chaque axe puis additionnés. Chaque convolution est précédée d'un suréchantillonnage d'un facteur 2, opéré en intercalant des zéros entre chaque échantillon.

borné $f(t)$ échantillonné en N éléments de résolution est une analyse multirésolution tronquée entre les approximations à la résolution du signal L et une résolution minimale S . Le signal est alors transformé en une somme de détails D_{2^j} , aux résolutions successives $J < j \leq L$ et de l'approximation A_{2^J} à la plus basse résolution, appelée “dernier lissé”, qui contient le résidu non-filtré de l'analyse. Les signaux détails étant de moyenne nulle, c'est le dernier lissé qui contient la normalisation du flux du signal:

$$\begin{aligned} f &= \sum_{j=-J}^L \langle f, \psi_{j,n} \rangle \psi_{j,n} + \langle f, \phi_{J,n} \rangle \phi_{J,n} \\ &= \sum_{j=-J}^L D_{2^j} + A_{2^J}. \end{aligned} \quad (2.44)$$

L'orthogonalité des bases d'ondelettes implique que les détails calculés aux différentes résolutions ne sont pas redondants mais complémentaires. La perte d'information entre les détails aux résolutions successives, due à la diminution de la fréquence de coupure de l'intervalle en fréquences, est compensée par une décimation d'un facteur 2 à chaque échelle. En outre, la dualité entre le signal et sa transformée se traduit par la conservation du nombre d'échantillons. On utilisera dans cette thèse des transformées en ondelettes d'images, c'est-à-dire de signaux à deux dimensions. La transformée d'une image $I(k,l)$ à deux dimensions ($\mathbf{N}_1 \times \mathbf{N}_2$) obtenue par l'algorithme multirésolution est généralement représentée par une image de même taille. Les signaux de détails aux approximations j successives sont des images de dimension $(\frac{\mathbf{N}_1}{2^j} \times \frac{\mathbf{N}_2}{2^j})$. On remplit l'image de la transformée en plaçant les trois images de signaux de détails à la résolution 2^0 , de dimension $\frac{\mathbf{N}_1}{2} \times \frac{\mathbf{N}_2}{2}$, aux trois coins de l'image. L'espace restant, de dimension $(\frac{\mathbf{N}_1}{2} \times \frac{\mathbf{N}_2}{2})$, est rempli dans les coins par les trois signaux de détails de dimension $(\frac{\mathbf{N}_1}{4} \times \frac{\mathbf{N}_2}{4})$, et ainsi de suite jusqu'à la résolution S , où le signal d'approximations associé au dernier lissé remplit l'espace restant (cf. 2.5).

2.3.2 Algorithme à trous.

Dans l'algorithme multirésolution, le sous-échantillonnage des signaux de détail par rapport au signal initial rend l'algorithme non redondant mais aussi non-invariant par translation. Si l'absence de redondance est par exemple recherchée en compression d'images, la redondance et l'invariance par translation sont au contraire utiles pour estimer les paramètres du signal.

L'algorithme à trous est un algorithme équivalent à un algorithme multirésolution sans décimations. En contrepartie la base d'ondelettes associée n'est pas orthogonale. Dans cet algorithme, les si-

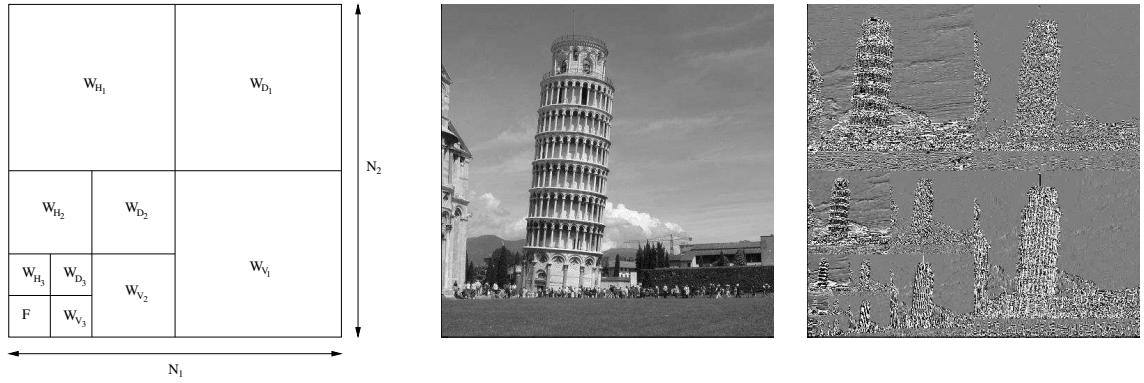


FIG. 2.5 – Diagramme de gauche: représentation de la transformée en ondelettes sur trois échelles d'une image de dimension $(N_1 \times N_2)$ obtenue par l'algorithme multirésolution à deux dimensions. A chaque résolution, la transformée comprend trois signaux de détails, W_H , W_V et W_D , respectivement horizontaux, verticaux et diagonaux. Le résidu à la résolution 3 est occupé par le dernier lissé F . Images de droite: Transformée en ondelettes de Haar obtenue selon l'algorithme multirésolution.

gnaux d'approximation ont tous le même nombre d'échantillons que le signal, ce qui rend l'algorithme invariant par translation et redondant. Comme pour l'algorithme multirésolution, les signaux d'approximations sont calculés par filtrages passe-bas successifs du signal $f(t)$ et les filtres d'analyse vérifient l'équation de dilatation (2.19). Par contre, puisque l'échantillonnage des signaux d'approximation est maintenant invariant d'échelle, c'est le filtrage qui ne l'est plus: l'équation (2.19) est vérifiée en définissant à chaque résolution j un filtre passe-bas h_j obtenu en insérant $2^j - 1$ zéros entre chaque échantillon du filtre h associé à l'analyse multirésolution, soit:

$$h_j[k]_{k \in \mathbb{Z}} = \begin{cases} h[n] & \text{si } k = 2^j n \\ 0 & \text{sinon.} \end{cases} \quad (2.45)$$

Par exemple, la décomposition de la figure (2.6), associée à un filtre \bar{h} à deux échantillons, fait intervenir des filtres \bar{h}_j à 2^{j+1} à 2^{j+1} échantillons dont 2^{j-1} zéros.

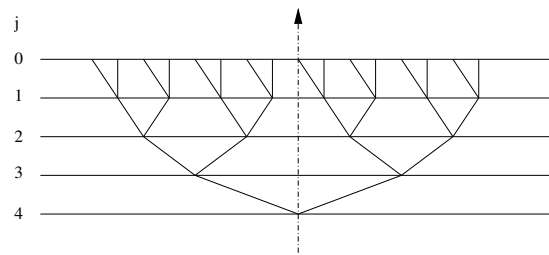


FIG. 2.6 – Décomposition d'un signal sur 4 échelles selon l'algorithme de Coifman & Donoho, avec un filtre d'analyse \bar{h}_j à deux échantillons non nuls, de taille 2^{j+1} .

La condition d'orthogonalité de la base d'ondelettes étant levée, on peut construire plusieurs familles d'ondelettes associées aux sous-espaces de détails. Une solution simple consiste à utiliser l'ondelette associée au filtre $g_j = 1 - h_j$, de sorte que le signal de détail W_j se construit à chaque résolution comme la différence des signaux d'approximations F_{j-1} et F_j (Holschneider). Cet algorithme est intuitif parce que la perte d'information entre les signaux d'approximations successifs se traduit par une différence, mais aussi rapide parce que la décomposition en signaux de détails, ainsi que la reconstruction du signal ne nécessitent pas de filtrage.

Une autre solution proposée par Coifman & Donoho (1995) consiste à construire le filtre passe-haut g_j de la même façon que le filtre h_j , en insérant $2^j - 1$ zéros entre chaque échantillon du filtre g associé à l'analyse multirésolution, soit :

$$g_j[k] = \begin{cases} g[n] & \text{si } k = 2^j n \\ 0 & \text{sinon.} \end{cases} \quad (2.46)$$

L'algorithme de Coifman & Donoho et l'algorithme à trous de Holschneider diffèrent en ce qui concerne la reconstruction du signal. En effet le filtrage de reconstruction de l'algorithme de Coifman & Donoho est aussi redondant que le filtrage de décomposition. En fait, ce filtrage revient à moyenner à chaque résolution les coefficients redondants de la décomposition. Par exemple pour un filtre \bar{g}_j à deux échantillons on effectue 2^j opérations à chaque échelle soit au total $2 * (2^J - 1)$ opérations au lieu de J dans le cas de l'algorithme à trous standard (cf. figure 2.6). Cette propriété rend l'algorithme intéressant pour réduire l'influence des fausses détections quand les coefficients sont modifiés ou seuillés.

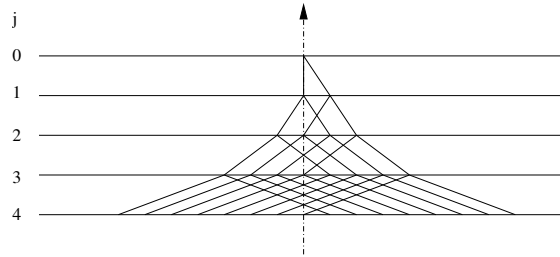


FIG. 2.7 – Reconstruction d'un signal sur 4 échelles selon l'algorithme de Coifman & Donoho, avec un filtre d'analyse \bar{g}_j à deux échantillons. A chaque échelle, 2^j coefficients redondants sont moyennés. Au total, le coefficient à l'échelle 0 est obtenu par $2 * (2^J - 1)$ opérations au lieu de J dans l'algorithme à trous standard.

En outre les deux algorithmes se distinguent dans le cas de l'analyse à deux dimensions. Considérons en effet l'analyse d'une image $I(k,l)$ de dimension $(N_1 \times N_2)$. L'algorithme à trous de Holschneider associe à chaque signal d'approximation F_j de l'image, un signal de détails de dimension $(N_1 \times N_2)$, $W_j = F_{j-1} - F_j$. Par contre l'algorithme de Coifman & Donoho qui utilise le formalisme de filtrage de l'analyse multirésolution à deux dimensions, associe trois sous-espaces de détails, respectivement, horizontaux, verticaux et diagonaux à chaque approximation F_j . L'algorithme de Coifman & Donoho à deux dimension est donc trois fois plus redondant que l'algorithme à trous de Holschneider et permet une exploration anisotrope des structures du signal.

2.3.3 Opérateurs de filtrage

Les deux algorithmes de calcul de transformées en ondelettes utilisés dans cette thèse sont l'algorithme multirésolution et l'algorithme de Coifman & Donoho. Ces deux algorithmes utilisent des relations de filtrages inter-échelles de décomposition et de reconstruction du type (2.37) et (2.39) à une dimension, (2.41) et (2.43) à deux dimensions. Pour l'algorithme multirésolution, non-redondant, ces relations sont invariantes d'échelles et font intervenir les filtres h et g . Pour l'algorithme de Coifman & Donoho, redondant, elles sont dépendantes de l'échelle et font intervenir les filtres h_j et g_j . Par la suite, on utilisera une écriture synthétique des relations de filtrage, faisant intervenir des opérateurs de filtrage.

A une dimension, on exprime la décomposition ou analyse du signal d'approximation F_j en signaux de détails W_{j+1} et d'approximations F_{j+1} par les relations :

$$\begin{aligned} F_{j+1} &= H_j F_j, \\ W_{j+1} &= G_j F_j. \end{aligned} \quad (2.47)$$

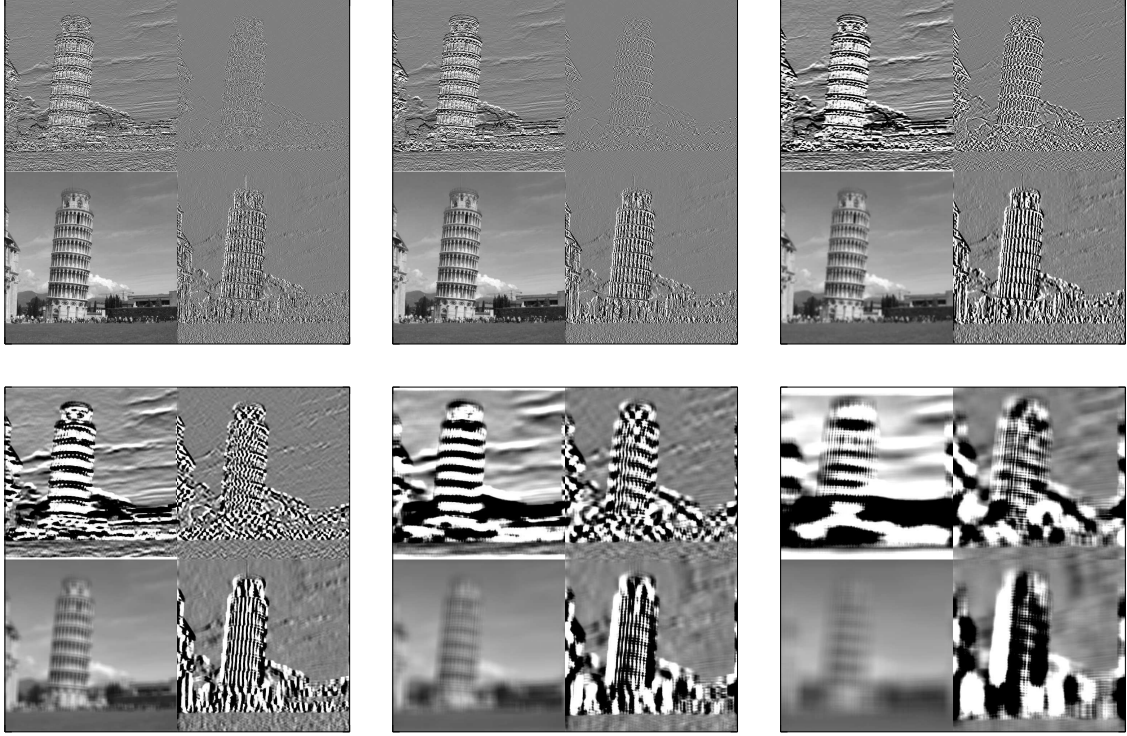


FIG. 2.8 – Transformée en ondelettes de Haar sur 6 échelles obtenue avec l'algorithme redondant de Coifman & Donoho. On a représenté pour chaque résolution le signal d'approximation et les trois signaux de détails de dimension $(N_1 \times N_2)$.

Dans le cas de l'algorithme multirésolution, les opérateurs H_j et G_j expriment la combinaison d'un filtrage par les filtres \bar{h} et \bar{g} respectivement et d'une décimation d'un facteur 2 (cf. figure (2.3)). Dans le cas de l'algorithme de Coifman & Donoho, ils expriment simplement les filtrages par h_j et \bar{g}_j .

La reconstruction ou synthèse du signal d'approximation F_j s'exprime quant à elle en fonction de W_{j+1} et F_{j+1} par:

$$F_j = \tilde{H}_j F_{j+1} + \tilde{G}_j W_{j+1}, \quad (2.48)$$

où les opérateurs \tilde{H}_j et \tilde{G}_j sont associés, soit à la combinaison d'un suréchantillonnage d'un facteur 2 et d'une convolution par h et g respectivement dans le cas de l'algorithme multirésolution (cf. figure (2.3), soit à une convolution par h_j et g_j dans le cas de l'algorithme de Coifman & Donoho².

A deux dimensions, le signal d'approximation F_j est décomposé en trois signaux de détail $W_{h,v,d}(j+1)$ et un signal d'approximation F_{j+1} , auxquels on associe quatre opérateurs d'analyse:

$$\begin{aligned} F(j+1) &= H_j F(j), \\ W_h(j+1) &= G_{h,j} F(j), \\ W_v(j+1) &= G_{v,j} F(j), \\ W_d(j+1) &= G_{d,j} F(j). \end{aligned} \quad (2.49)$$

2. L'introduction des notations H_j et G_j d'une part, \tilde{H}_j et \tilde{G}_j d'autre part, est introduite pour sa cohérence avec l'analyse en ondelettes bi-orthogonales, mais nous n'utiliserons pas d'ondelettes bi-orthogonales dans cette thèse

Les opérateurs H_j , $G_{h,j}$, $G_{v,j}$, $G_{d,j}$ sont alors associés, soit aux combinaisons de filtrages décimés par $\bar{h}\bar{h}$, $\bar{h}\bar{g}$, $\bar{g}\bar{h}$ et $\bar{g}\bar{g}$ dans le cas de l'algorithme multirésolution, soit aux filtrages par $\bar{h}_j\bar{h}_j$, $\bar{h}_j\bar{g}_j$, $\bar{g}_j\bar{h}_j$ et $\bar{g}_j\bar{g}_j$ dans le cas de l'algorithme de Coifman & Donoho.

Enfin la reconstruction de F_j s'exprime par:

$$F_j = \tilde{H}_j F(j+1) + \tilde{G}_{h,j} W_h(j+1) + \tilde{G}_{v,j} W_v(j+1) + \tilde{G}_{d,j} W_d(j+1), \quad (2.50)$$

et les opérateurs \tilde{H}_j , $\tilde{G}_{h,j}$, $\tilde{G}_{v,j}$ et $\tilde{G}_{d,j}$ sont associés, soit aux combinaisons de filtrages suréchantillonnés hh , hg , gh et gg dans le cas de l'algorithme multirésolution, soit aux filtrages par $h_j h_j$, $h_j g_j$, $g_j h_j$ et $g_j g_j$ dans le cas de l'algorithme de Coifman & Donoho.

2.4 Applications: Compression de l'information et débruitage.

En décomposant un signal temporel en ses variations dans l'espace temps-fréquence, la transformée en ondelettes permet de coder l'information contenue dans le signal dans un petit nombre de coefficients de grande amplitude. Cette propriété la rend particulièrement efficace pour deux applications: la compression de l'information et le débruitage.

Pour la compression, on applique aux coefficients de la transformée un opérateur de quantification Q , en conséquence tous les coefficients de faible amplitude, qui sont peu porteurs d'information, sont par exemple annulés:

$$\tilde{f} = \sum_{j=-J}^L Q(\langle f, \psi_{j,n} \rangle) \psi_{j,n} + \langle f, \phi_{J,n} \rangle \phi_{J,n}. \quad (2.51)$$

En utilisant une transformée non redondante comme l'algorithme multirésolution, le fait d'annuler un grand nombre de coefficients ramène le nombre d'échantillons non-nuls à un nombre beaucoup plus faible que le nombre d'échantillons du signal de départ.

Pour le débruitage, on utilise cette fois l'espace temps-fréquence pour séparer les contributions signal et du bruit. Le bruit étant supposé plus lentement variable que le signal à chaque échelle, les coefficients de sa transformée sont de faible amplitude et non localisés. Inversement les coefficients associés aux transitoires du signal sont d'amplitude forte et localisés. Connaissant la nature du bruit, on peut en modéliser puis en supprimer la contribution uniforme dans l'espace des ondelettes, en appliquant des contraintes à l'ensemble des coefficients. Ces contraintes d , peuvent être linéaires, -filtrage-, ou non linéaires, -seuillage-:

$$\tilde{f} = \sum_{j=-J}^L d(\langle f, \psi_{j,n} \rangle) \psi_{j,n} + \langle f, \phi_{J,n} \rangle \phi_{J,n}. \quad (2.52)$$

Les contraintes de seuillage introduisent des discontinuités dans l'espace temps-fréquence qui se traduisent par des oscillations dans l'espace direct. Ces artefacts peuvent être réduits en utilisant cette fois des transformées redondantes dans lesquelles les oscillations de même fréquence mais de phases distinctes seront moyennées et annulées. C'est le débruitage qui justifie l'utilisation de la transformée en ondelettes dans cette thèse, son principe sera développée dans la section 3.2.1.

2.5 Choix d'une base d'ondelettes orthogonales.

En définissant des sous-espaces duaux de détails et d'approximations, l'analyse multirésolution permet de définir des bases d'ondelettes orthogonales à partir de bases d'interpolation ou fonctions d'échelle. En utilisant (2.21) et (2.22) on construit ainsi dans l'espace de Fourier une ondelette orthogonale en fonction d'une fonction d'échelle ϕ et du filtre d'approximation h associé:

$$\hat{\psi}[\omega] = \frac{1}{\sqrt{2}} e^{\frac{-i\omega}{2}} \hat{h}^* \left[\frac{\omega}{2} + \pi \right] \hat{\phi} \frac{\omega}{2}. \quad (2.53)$$

La construction la plus “triviale” est obtenue à partir d’une analyse multirésolution associée à une fonction d’échelle constante par morceaux, $\phi = \mathbf{1}_{[0,1]}$. Son filtre d’analyse à deux échantillons se déduit de l’équation de dilatation (2.19):

$$\begin{aligned} h[n] &= \left\langle \frac{1}{\sqrt{2}} \phi\left(\frac{t}{2}\right), \phi(t-n) \right\rangle \\ &= \begin{cases} \frac{1}{\sqrt{2}} & \text{si } n = 0, 1 \\ 0 & \text{sinon.} \end{cases} \end{aligned} \quad (2.54)$$

Il s’agit de l’ondelette de Haar, qui s’exprime en fonction de h et ϕ en utilisant les relations (2.20) et (2.23):

$$\begin{aligned} \frac{1}{\sqrt{2}} \psi\left(\frac{t}{2}\right) &= \sum_{n=-\infty}^{+\infty} (-1)^{1-n} h[1-n] \phi(t-n), \\ &= \frac{1}{\sqrt{2}} [\phi(t-1) - \phi(t)] .. \end{aligned} \quad (2.55)$$

On en déduit que l’ondelette de Haar est une fonction à deux valeurs, 1 et -1, (Haar, 1910) de support $[0,1]$ (cf. figure 2.9):

$$\psi(t) = \begin{cases} -1 & \text{si } 0 \leq t < 1/2 \\ 1 & \text{si } 1/2 \leq t < 1 \\ 0 & \text{sinon} \end{cases} \quad (2.56)$$

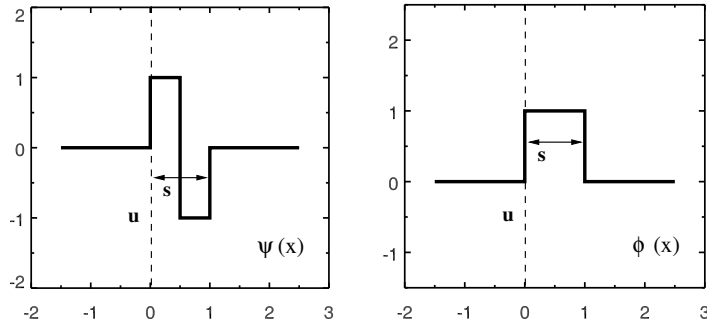


FIG. 2.9 – Ondelette de Haar ψ et fonction d’échelle indicatrice ϕ associée.

En utilisant des bases d’interpolations plus régulières que la fonction indicatrice, on peut former d’autres ondelettes. Par exemple les ondelettes splines polynômiales sont formées à partir d’approximations multirésolution par splines.

On a vu que l’intérêt de la transformée en ondelettes est de coder l’information contenue dans le signal dans un nombre minimal de coefficients de grande amplitude. Cet objectif sera d’autant mieux atteint si on choisit une ondelette dont la forme est, d’une part, corrélée avec les régions rapidement variables du signal f qu’elle analyse, d’autre part décorrélée avec ses régions plus régulières.

Dans les régions où le signal présente de fortes discontinuités, les coefficients en ondelettes associés à ces discontinuités seront de grande amplitude à toutes les échelles et c’est la taille du support de l’ondelette qui en limitera le nombre dans l’ensemble de la transformée. Dans les régions plus régulières du signal, on pourra réduire l’amplitude des coefficients en augmentant le nombre de moments nuls $\{\mu_k\}_{0 \leq k < p}$ de l’ondelette:

$$\mu_k = \int_{-\infty}^{+\infty} t^k \psi(t) dt = 0 \quad 0 \leq k < p. \quad (2.57)$$

On déduit en effet de cette expression qu'une ondelette dont les p premiers moments sont nuls est orthogonale aux différents termes d'un polynôme d'ordre p par lequel on peut approcher le signal.

Le choix d'une ondelette dépend donc du signal analysé. C'est un compromis entre la minimisation de la taille du support pour les signaux présentant de fortes discontinuités et la maximisation du nombre de moments nuls pour les signaux plus réguliers. Les ondelettes de Daubechies (1988) se construisent par exemple numériquement en minimisant la taille du support pour un nombre de moments non nul donné.

L'ondelette de Haar est discontinue et ne possède qu'un seul moment non nul. Si son support est le plus compact possible, -c'est une ondelette de Daubechies à un moment non nul-, elle est par contre mal adaptée à l'analyse de signaux continus. Nous avons choisi de l'utiliser dans cette thèse pour tirer parti de la simplicité de son expression numérique, mais surtout des propriétés particulières de l'analyse multirésolution associée à une fonction d'échelle constante par morceaux. Nous utiliserons ainsi des algorithmes de débruitage basés sur le seuillage de la transformée en ondelettes de Haar. N'utilisant pas d'ondelette plus régulière, ces processus de seuillage entraînent de fortes discontinuités que nous avons choisi de réduire en introduisant une contrainte de "régularisation" du signal dans les algorithmes de reconstruction du signal.

Chapitre 3

Restauration d'image à faible statistique appliquée à l'émission X du plasma intra-amas.

3.1 Le signal: Emissivité du plasma intra-amas.

Disposant d'une liste de N événements (k, l, e) observés, nous cherchons à restituer de manière optimale la distribution d'émissivité du plasma intra-amas projetée sur le plan du ciel, $S(k, l) = \int_0^\infty S(e, k, l) de$. Il s'agit d'un problème inverse que l'on posera après avoir rappelé pourquoi le processus d'émission de l'ensemble des photons reçus sur le plan du détecteur est assimilable à un processus de Poisson.

3.1.1 Du processus de Bernoulli au processus de Poisson

Pour examiner le processus discret d'émission de l'ensemble des photons reçus, échantillons l'espace des événements en éléments de résolution de position (k, l) et énergie (e) . L'ensemble des N événements étant associé à différentes composantes d'émission, associons une loi de probabilité $f_i(k, l, e)$ et une émissivité totale n_i à chaque composante i , ainsi qu'une loi de probabilité globale $F(k, l, e)$ à l'ensemble des événements:

$$N_F F(k, l, e) = \sum_i n_i f_i(k, l, e) \quad (3.1)$$

Par exemple, pour les spectro-imageurs EPIC, on utilisera un modèle à trois composantes, comprenant une composante source et deux composantes de fond de lois de probabilités respectives, $S(k, l, e)$, $P(e)$, $B(e)$, et d'émissivités totales N_S , N_P , N_B (cf. eq. 1.2). Pour chaque élément de résolution k, l, e , cherchons le nombre de photons émis par composante du signal, connaissant la probabilité d'émission d'un photon, $f_i(k, l, e)$ et le nombre total de photons émis, n_i . Il s'agit d'un ensemble de N expériences de Bernoulli de paramètre $f_i(k, l, e)$. Le nombre de photons émis par composante suit donc la variable aléatoire binomiale $B(n_i, p_i)$, avec $p_i = f_i(k, l, e)$, et la probabilité d'observer k photons s'exprime par:

$$P(X_i = k) = \frac{n_i!}{k!(n_i - k)!} p_i^k (1 - p_i)^{n_i - k}. \quad (3.2)$$

Cet échantillonnage de l'espace des événements conduisant à de faibles probabilités d'émission par élément de résolution, $f_i(k, l, e) \ll 1$ et un grand nombre de photons étant émis durant l'exposition, on peut approcher la variable aléatoire Binomiale $B(n_i, p_i)$ par une loi de Poisson $P(\lambda_i)$ de paramètre $\lambda_i = n_i p_i$. La probabilité d'observer k photons par élément de résolution et par composante du signal s'exprime alors par:

$$P(X_i = k) = \exp(-\lambda_i) \frac{\lambda_i^k}{k!}. \quad (3.3)$$

Finalement le nombre total de photons reçus dans l'élément de résolution (k, l, e) suit une variable aléatoire qui est la somme des lois de Poisson de paramètre λ_i associées à chacune des composantes du signal. Puisque la somme de deux variables aléatoires de Poisson de paramètres respectifs λ_1 et λ_2 est une loi de Poisson de paramètre $\lambda_1 + \lambda_2$, on en déduit de proche en proche que cette variable aléatoire est une loi de Poisson de paramètre $\lambda = \sum_i \lambda_i$.

3.1.2 Une source étendue avec des variations spatiales.

Ayant modélisé la probabilité d'émission des photons dans l'espace des événements par la loi $F(k, l, e)$, on suppose connue la distribution spectrale de chacune des composantes d'émission $f_i(e)$, mais on cherche à restituer leurs distributions spatiales. Dans le cas général, on ne peut étudier séparément les variations spatiales de chacune des composantes, puisque les événements associés à chacune d'entre elles sont indiscernables. Heureusement, le problème est ici simplifié puisque notre modèle comprend une composante associée au milieu intra-amas $S(k, l, e)$, dont on cherche les variations spatiales $S(k, l) = \sum_e S(k, l, e)$ et deux composantes de fond, P et B , dont les distributions spatiales sont supposées uniformes. L'étude des variations spatiales de la distribution

des photons du milieu intra-amas $S(k,l)$ se ramène donc à l'étude des variations spatiales de la distribution de l'ensemble des photons $F(k,l) = \sum_e F(k,l,e)$.

L'espace des paramètres à trois dimensions (k,l,e) est maintenant contracté en un espace à deux dimensions (k,l) et le nombre total de photons reçus dans l'élément de résolution (k,l) suit une variable aléatoire qui est la somme des lois de Poisson de paramètre $\lambda = F(k,l,e)$ associées à l'émissivité de chaque élément de résolution (k,l,e) . C'est donc encore une loi de Poisson avec cette fois pour paramètre $\lambda(k,l) = \sum_e \lambda(k,l,e)$. Le processus d'émission de l'ensemble des N photons reçus sur le plan du détecteur est donc assimilable à un processus de Poisson de paramètre $\lambda(k,l)$:

$$P[X(k,l) = m] = \exp(-\lambda(k,l)) \frac{\lambda(k,l)^m}{m!}. \quad (3.4)$$

Et le problème de l'étude des variations spatiales d'émissivité du plasma intra-amas se ramène à l'étude des variations spatiales du paramètre de Poisson, $\lambda(k,l)$.

3.2 Choix de l'algorithme

Dans la section précédente, nous avons ramené le problème de l'étude des variations spatiales d'émissivité du milieu intra-amas à celui de l'estimation de la distribution spatiale d'un paramètre de Poisson $\lambda(k,l)$ associé au processus d'émission des N photons observés au foyer des spectro-imageurs EPIC. La distribution spatiale du paramètre $\lambda(k,l)$ est issue de processus physiques déterministes liés au milieu intra-amas, milieu continu dans un espace à trois dimensions. Le processus d'émission des photons observés sur le plan du détecteur n'est donc pas stationnaire, par contre la continuité du milieu intra-amas permet de le considérer comme localement stationnaire. Cette hypothèse de stationnarité locale permet de définir des éléments de résolution spatiaux dans lesquels on peut estimer le paramètre λ . Pour les définir sans a priori, il est intéressant d'explorer au préalable de manière systématique le signal et ses variations à différentes résolutions spatiales en utilisant un algorithme adaptatif. Nous discuterons des avantages et inconvénients de plusieurs algorithmes multirésolution adaptés à notre problème dans la section 3.2.1, avant de détailler celui que nous avons choisi dans la partie 3.3 de ce chapitre. L'estimation locale du paramètre de Poisson proprement dit fera l'objet de la section 3.2.2.

3.2.1 L'approche multiéchelle

Ayant choisi d'explorer le signal dans le cadre d'une approche multiéchelle, on pourra utiliser deux types d'algorithmes, les algorithmes à maillage adaptatif ou la transformation en ondelettes. Dans les algorithmes à maillage adaptatif, l'espace est directement rééchantillonné en éléments de résolution de tailles distinctes. Avec la transformation en ondelettes, on effectue une exploration de l'espace position-fréquence en codant par des coefficients les variations spatiales du signal à différentes échelles.

Les algorithmes à maillage adaptatif

Le principe des algorithmes à maillage adaptatif est de convoluer la distribution de points par un noyau de forme prédéfinie mais de taille ou échelle variable. L'échelle du noyau doit être minimisée pour préserver la résolution spatiale mais suffisante pour garantir un rapport signal-sur-bruit supérieur à une valeur seuil. Ces algorithmes se distinguent par la forme du noyau, le caractère compact ou non de son support, mais aussi l'échantillonnage de l'espace position-fréquence qu'ils sous-tendent. Ils sont très utilisés pour l'étude des amas de galaxies, à la fois en astronomie X pour caractériser la distribution d'émissivité du milieu intra-amas, mais aussi dans le domaine optique pour étudier des distributions de densité de galaxies, envisagées comme des points.

En astronomie X, deux algorithmes sont couramment utilisés, l'échantillonnage adaptatif et le lissage adaptatif (Ebeling et al., 1998). Dans l'échantillonnage adaptatif, la densité d'événements est d'abord évaluée dans chaque pixel de l'image, puis la surface des éléments de résolution est

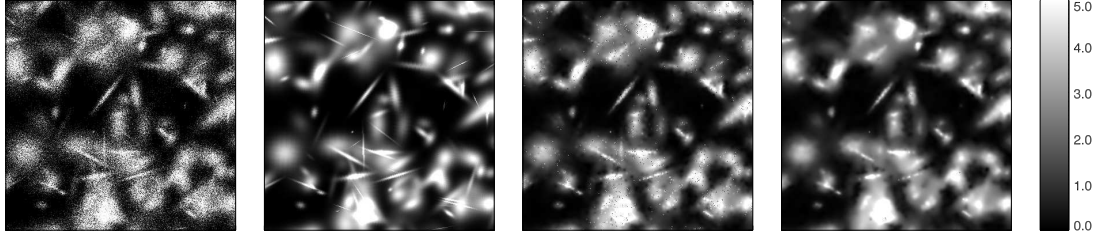


FIG. 3.1 – *Lissage adaptatif gaussien. (1) Fluctuation de Poisson (5 dB). (2) Image modèle. (3 et 4) Lissage adaptatif à 2 et 3 σ (14.2 dB et 17.4 dB).*

progressivement augmentée en groupant les pixels quatre par quatre jusqu'à ce que le rapport signal-sur-bruit dépasse la valeur seuil. Dans le lissage adaptatif, la distribution de points est convoluée par des noyaux d'échelle croissante, la valeur de densité locale retenue étant celle calculée avec le noyau d'échelle minimale pour que le rapport signal-sur-bruit dépasse la valeur seuil. Ces deux algorithmes se distinguent par le fait que l'échantillonnage adaptatif n'est pas redondant et utilise un échantillonnage dyadique des échelles de calcul. Il nécessite donc beaucoup moins d'opérations que le lissage adaptatif, mais génère plus d'artefacts. En outre le lissage adaptatif permet d'implémenter des noyaux de forme variée.

En astronomie optique, A. Dressler (1980) a travaillé sur la structuration de la distribution projetée des galaxies dans des échantillons d'amas. Dans les amas relaxés, on montre en appliquant le théorème du Viriel que cette distribution suit une loi de Poisson dont le paramètre local suit lui-même une loi de type modèle β (cf. équation (2.2) de la partie I). La mise en évidence de structurations dans la distribution des galaxies est donc un indicateur de leur état de relaxation. Pour ses travaux, A. Dressler a proposé un algorithme dans lequel la densité locale de galaxies est caractérisée par l'inverse de la surface minimale incluant un nombre prédéfini de galaxies (Kriessler, 1997).

Dans ces trois algorithmes l'échelle du noyau de convolution est optimisée en fonction du rapport signal-sur-bruit. Dans les algorithmes d'échantillonnage et de lissage adaptatif, une distribution de fond est prise en compte. Le signal est le nombre d'événements dépassant la valeur de fond. Le "bruit" est calculé selon une statistique gaussienne en sommant les erreurs quadratiques associées au calcul de la densité locale de photons associés à l'amas et au fond. Dans l'algorithme de Dressler, on suppose une distribution poissonnienne des événements et il n'y a pas de distribution de fond. La fluctuation du bruit se ramène donc à la racine du nombre d'événements, et peut être seuillée par un simple nombre d'événements. En fait l'algorithme de Dressler se ramène à un lissage adaptatif sans distribution de fond et avec un noyau uniforme.

La figure 3.1 illustre le lissage adaptatif gaussien d'une distribution de points (1) dont on cherche à restaurer le paramètre de Poisson (2). Le seuil de signification du signal est fixé respectivement à 2 et 3 fois l'écart-type du bruit σ pour les images (3) et (4). On constate qu'avec un critère de signification conservatif de 3 σ , l'algorithme restaure de manière satisfaisante les structures à grande échelle. Si on diminue le rapport signal-sur-bruit, on fait apparaître de petites structures dans les régions où le paramètre de Poisson local varie rapidement, mais la forme de ces petits objets est assez bruitée et distordue par rapport au modèle. Nous avons vu que l'estimation du paramètre de Poisson devait être effectuée dans des domaines de l'espace temps-fréquence où le signal est localement stationnaire. Les algorithmes à maillage adaptatif n'atteignent qu'approximativement cet objectif en adaptant l'échantillonnage pour homogénéiser le rapport signal-sur bruit et non pour rendre le signal localement stationnaire. Par ailleurs, ils n'explorent que partiellement l'espace temps-fréquence puisqu'une échelle de convolution unique est attribuée à chaque pixel de l'image, ce qui ne permet pas de restituer la nature multiéchelle des objets d'une image telle que celle de la figure 3.1. La transformée en ondelettes permet de résoudre ces problèmes en explorant de manière plus complète l'espace temps-fréquence et en donnant directement une mesure de la stationnarité du signal. Elle permet aussi de découpler l'analyse temps-fréquence du débruitage et

ainsi de construire des algorithmes itératifs plus performants.

Les algorithmes ondelettes et la séparation signal-bruit

Avec la transformée en ondelettes, les variations du signal $S(k,l)$ à différentes échelles a sont codées par des coefficients. Cette approche multi-échelle "naturelle" permet de coder l'essentiel de l'énergie du signal dans un petit nombre de coefficients localisés. Comme la distribution spatiale du bruit est en général plus faiblement variable que celle du signal, les coefficients associés au bruit sont distribués de manière plus continue que ceux associés au signal et leur amplitude est plus faible. C'est pourquoi l'analyse en ondelettes permet de séparer efficacement les contributions respectives du bruit et du signal dans l'espace position-fréquence.

La figure 3.2 illustre le principe de la séparation signal-bruit dans l'espace des ondelettes. On y a représenté trois signaux à une dimension, ainsi que leurs transformées en ondelettes respectives calculées sur six échelles en utilisant une ondelette de Haar. Le signal du haut est composé de la somme d'une gaussienne large et d'une gaussienne étroite. Le signal du bas à gauche est un bruit blanc. Le signal du bas à droite est un signal bruité qui est la superposition des deux précédents. On constate dans l'espace des ondelettes, que l'amplitude des coefficients est maximale aux résolutions associées aux largeurs caractéristiques de chacun des objets. Ainsi l'amplitude des coefficients associés à la gaussienne large est maximale à la 6^e échelle, alors que l'amplitude des coefficients associés à la gaussienne étroite est maximale à la 3^e échelle. Enfin, comme la taille caractéristique du bruit blanc est l'échantillonnage du détecteur, l'amplitude des coefficients associés au bruit blanc est maximale à la première échelle et décroît ensuite avec l'échelle.

En fait, comme la transformée en ondelettes est linéaire, l'amplitude des coefficients associés au bruit est modélisable si on connaît la nature du bruit. On s'intéressera par la suite à deux types de signaux: d'une part, les signaux de type $X = F + W$, bruités par un bruit blanc, additif et gaussien, $W(\sigma)$, d'écart-type σ , d'autre part les processus de Poisson de type $X = p_\lambda(n)$.

Pour un signal bruité par un bruit gaussien, on montre que la densité de probabilité des coefficients en ondelettes liés au bruit est elle aussi une gaussienne de déviation standard σ et d'amplitude dépendant de la normalisation en énergie des coefficients. Pour un processus de Poisson, on ne peut séparer les fluctuations du bruit et du signal, qui dépendent toutes deux du paramètre de Poisson λ . Si la densité d'événements n'est pas trop faible ($n > 6$), la transformée d'Anscombe $t(X) = 2\sqrt{X + \frac{3}{8}}$ permet de stabiliser la variance du signal $X = p_\lambda(n)$ et de se ramener au cas du bruit additif.

Si la densité d'événements est plus faible, on doit estimer la fluctuation du bruit dans l'espace des ondelettes en calculant la densité de probabilité des coefficients en ondelettes associés à un processus de Poisson stationnaire. Pour cela on peut d'abord exprimer la densité de probabilité des coefficients associés à n événements distribués aléatoirement dans l'intervalle couvert par le support de l'ondelette ψ . Considérant que la distribution d'un événement dans l'espace des ondelettes est directement donnée par l'histogramme H_ψ de l'ondelette ψ , on en déduit, en exprimant de proche en proche la loi de probabilité des couples $(P(w|i), P(w|1), i = 1, \dots, n)$ que la distribution des n événements indépendants est donnée par n autoconvolutions de H_ψ (Starck & al, 1998):

$$\begin{aligned} p(w|n) &= H_\psi \otimes H_\psi \otimes \dots \otimes H_\psi \\ &= H_\psi^{(*n)}(w) \end{aligned} \quad (3.5)$$

La distribution de probabilité des coefficients en ondelettes associés à un processus de Poisson stationnaire $p_\lambda(n)$, de paramètre λ s'exprime alors, selon le théorème des probabilités totales, par:

$$\begin{aligned} p_w &= \sum_{n=0}^{\infty} p(w|n) p_\lambda(n) \\ &= e^{-\lambda} \delta(w) + \sum_{n=1}^{\infty} H_\psi^{(*n)}(w) \frac{\lambda^n}{n!} e^{-\lambda}, \end{aligned} \quad (3.6)$$

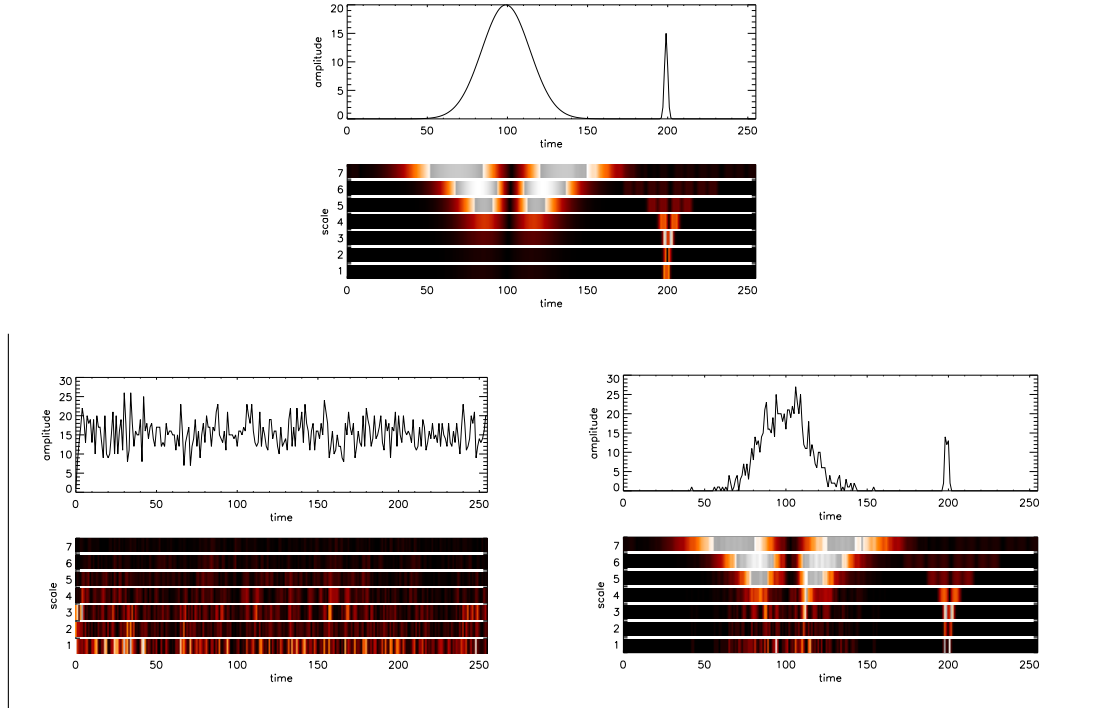


FIG. 3.2 – Séparation signal-bruit dans l'espace des ondelettes. Représentation de signaux et leurs transformées en ondelettes calculées sur 7 échelles en utilisant l'ondelette de Haar. Haut: Signal 1D. Bas-gauche: Bruit blanc. Bas-droite: Signal bruité.

où l'on a considéré que la densité de probabilité des coefficients en ondelettes se ramène à un dirac centré en 0 dans le cas particulier de l'absence d'événements, $n=0$.

3.2.2 Débruitage

Ayant projeté le signal $S(k,l)$ dans un espace multi-échelle, il reste à choisir la méthode de débruitage la plus appropriée pour restituer la distribution spatiale du paramètre $\lambda(k,l)$ à partir des coefficients en ondelettes $W(k,l)$ associés au signal bruité $S(k,l)$. Le signal F sera alors estimé à l'aide de l'estimateur \tilde{F} , en fonction des données bruitées X et d'un opérateur de décision D :

$$\tilde{F} = DX. \quad (3.7)$$

Considérant le signal cherché, F , comme la réalisation de la variable aléatoire π , on pourra envisager le débruitage selon deux approches. Si π est connue ou modélisable, on cherchera le signal F le plus probable en tenant compte du signal bruité observé, c'est l'approche bayésienne. Si π est inconnue on cherchera à estimer les fluctuations du bruit pour les supprimer du signal avec un niveau de confiance choisi, c'est l'approche fréquentiste.

Algorithmes Bayesiens.

L'estimation bayésienne repose sur l'estimateur \tilde{F} qui minimise le "risque", $r(D,\pi) = E\{\|F - \tilde{F}\|^2\}$. On montre qu'il s'agit de l'espérance conditionnelle de F compte-tenu des données bruitées, $\tilde{F} = E\{F|X\}$. On la détermine en général à partir de l'espérance conditionnelle des données compte-tenu de F , $E\{X|F\}$, en utilisant la formule de Bayes, $p(F|X) = \frac{p(F)p(X|F)}{p(X)}$. La distribution de probabilités π des signaux F n'étant pas toujours a priori connue, une variante de l'approche

bayésienne consiste à minimiser le risque $r(D, \pi)$ pour la distribution de probabilité la moins favorable des distributions a priori, c'est l'estimation minimax.

L'estimateur bayésien est souvent difficile à exprimer. On peut simplifier le problème en minimisant $r(D, \pi)$ dans l'espace des opérateurs de décision linéaires. L'estimateur linéaire \tilde{F} est alors appelé estimateur de Wiener. Dans le cas où les variables aléatoires F et W associées au signal et au bruit sont gaussiennes, on montre que le filtre de Wiener est optimal parmi les estimateurs linéaires et non linéaires de F, et qu'il constitue un estimateur bayésien. Si de plus F et W correspondent à des processus stationnaires, l'estimateur de Wiener est diagonal dans une base de Fourier discrète, $\{g_m\}_{m=0, \dots, n}$, qui diagonalise les matrices de covariance du signal et du bruit. Il s'exprime alors en fonction des valeurs propres de ces matrices, β_m^2 et σ_m^2 , aussi appelées puissances spectrales:

$$\tilde{F} = \sum_{m=0}^{N-1} \frac{\beta_m^2}{\beta_m^2 + \sigma_m^2} X g_m \quad (3.8)$$

Si l'on peut construire des bases de l'espace direct où β_m^2 et σ_m^2 sont invariantes, ce filtrage dans l'espace de Fourier se ramène à une convolution par le filtre de Wiener, $h = \frac{\beta_m^2}{\beta_m^2 + \sigma_m^2}$. Le filtre h se comporte alors comme un filtre passe-bas dépendant du rapport signal sur-bruit $\frac{\beta_m^2}{\sigma_m^2}$: il ne modifie pas le signal si $\sigma_m^2 \ll \beta_m^2$ mais l'atténue d'autant plus que $\sigma_m^2 \gg \beta_m^2$. Les bases d'ondelettes rendant invariantes β_m^2 et σ_m^2 à une échelle a donnée, on peut y construire un filtre de type Wiener associé à chaque échelle.

Filtrage de Wiener dans le cas d'un bruit additif gaussien. Starck & Bijaoui (1994) ont proposé un algorithme de filtrage dans l'espace des ondelettes, dans le cas où la variable aléatoire F est la réalisation d'un processus gaussien uniforme d'écart-type σ_F . Le filtre de type Wiener exprimé à chaque échelle a dans la base des ondelettes s'exprime comme dans les bases de Fourier par $\alpha_j = \frac{\sigma_F^2}{\sigma_F^2 + \sigma^2}$. L'hypothèse d'uniformité du signal étant en général trop forte, Bijaoui (2000) a proposé un algorithme qui s'applique à tous les signaux dont la variable aléatoire F est la réalisation d'une somme de processus gaussiens. Pour ces deux algorithmes, la linéarité de la transformée en ondelettes et la régularité du signal permettent d'améliorer de façon hiérarchique l'estimation des coefficients à partir de l'estimation des coefficients aux plus basses résolutions, en modifiant itérativement l'espérance conditionnelle de F.

Processus de Poisson Un algorithme de filtrage de type Wiener a été adapté au processus de Poisson par Nowak & Baraniuk (1997). Cet algorithme estime les puissances respectives du signal et du bruit des coefficients en ondelettes, $\sigma_{W_\lambda}^2$ et $\sigma_{W_C}^2$, à partir de leur histogramme. La transformée est alors filtrée par le filtre $h = \frac{\sigma_{W_C}^2}{\sigma_{W_C}^2 + \sigma_{W_\lambda}^2}$. Ce filtre se comporte en fait asymptotiquement avec le nombre d'événements N comme le filtrage de Wiener d'un bruit additif gaussien de variance $\sigma = \sqrt{N}$. Cette estimation est donc asymptotiquement optimale mais fautive dans le cas d'un faible nombre d'événements où la variance du filtre est alors surestimée et le filtrage trop conservatif.

La construction d'un véritable algorithme de filtrage bayésien adapté au processus de Poisson est complexe parce que l'hypothèse de stationnarité et la linéarité de la transformée en ondelettes imposent de construire des algorithmes itératifs et hiérarchiques. En effet, le paramètre de Poisson doit être estimé de proche en proche, à la fois dans les directions d'espace et l'axe des fréquences. L'algorithme de Timmermann & Nowak (1999) propose de résoudre le problème en partant d'une variable aléatoire locale a priori, f_Λ , dont le paramètre de Poisson λ est une réalisation.

Algorithmes fréquentistes: Seuillage des coefficients en ondelettes.

Algorithmes	Estimation du bruit	Itérativité	Redondance spatiale	Estimation du fond	Type
Filtrage ondelettes (Starck / Bijaoui)	puissance	oui	oui	non	bayésien
Seuillage ondelettes ($k\text{-}\sigma$ / Donoho)	PDF exacte	non	oui	non	fréquentiste

FIG. 3.3 – Algorithmes d'estimation du signal dans le cas d'un bruit additif.

Algorithmes	Estimation du bruit	Itérativité	Redondance spatiale	Estimation du fond	Type
Echantillonnage adaptatif (Ebeling)	variance	non	non	oui	fréquentiste
Lissage adaptatif (Ebeling)	variance	non	oui	oui	fréquentiste
Lissage adaptatif (Dressler)	variance	non	oui	non	fréquentiste
Filtrage ondelettes (Nowak & Baraniuk)	puissance	non	oui	non	fréquentiste
Seuillage ondelettes (Kolaczyk)	PDF approchée	non	oui	non	fréquentiste
Seuillage ondelettes (Jammal & Bijaoui)	PDF exacte	oui	oui	non	fréquentiste
Estimation ondelettes (Timmermann & Nowak)	variance	oui	oui	possible	bayésien

FIG. 3.4 – Algorithmes d'estimation du signal adaptés au processus de Poisson.

L'approche fréquentiste repose sur l'estimation de la fluctuation du bruit à partir des données. Elle est moins performante que l'approche bayésienne qui utilise une information sur la distribution de probabilités du signal cherché, mais elle est toujours utilisable dans les cas où on ne dispose pas de cette information a priori.

L'espace des ondelettes est particulièrement adapté à l'estimation du bruit parce qu'il permet de séparer efficacement le signal du bruit en concentrant l'énergie du signal sur quelques coefficients alors que le bruit est déployé sur l'ensemble de la transformée (cf. section 3.2.1). L'estimation de la fluctuation du bruit permet alors d'appliquer un seuil de signification T à l'ensemble des coefficients en ondelettes et de les séparer en coefficients significatifs si leur amplitude dépasse la valeur seuil et non significatifs sinon. En fait, dans le cas général, on définit une valeur seuil positive t_+ pour les coefficients en ondelettes positifs et une valeur seuil négative t_- pour les coefficients en ondelettes négatifs. Le signal F est alors estimé à partir sa transformée en ondelette "seuillée" par un projecteur non-linéaire D qui applique des contraintes distinctes $d_W(W)$ aux coefficients significatifs et non-significatifs. Notons que ce type d'algorithme est en général beaucoup plus rapide qu'un algorithme de lissage.

$$\tilde{F} = DW = \sum_{m=0}^{N-1} d_W(W)\psi_m \quad (3.9)$$

La contrainte la plus triviale est d'annuler les coefficients en ondelettes non-significatifs, c'est le "seuillage dur". Cette solution a l'avantage de ne pas modifier la partie significative du signal, mais l'inconvénient d'introduire des discontinuités dans la décomposition temps-fréquence, donc dans le signal reconstruit final.

$$d_{hard}(a,k,l) = \begin{cases} W(a,k,l) & \text{si } W(a,k,l) > t_+ \\ 0 & \text{si } t_- \leq W(a,k,l) \leq t_+ \\ W(a,k,l) & \text{si } W(a,k,l) < t_- \end{cases} \quad (3.10)$$

Le "seuillage mou" consiste à retrancher la valeur seuil positive t_+ à tous les coefficients positifs et à annuler les coefficients négatifs obtenus après la soustraction. Pour les coefficients négatifs, la valeur seuil négative t_- est au contraire ajoutée et les coefficients positifs obtenus annulés. Cette solution comme la précédente annule les coefficients du domaine non-significatif, mais réduit les discontinuités entre les coefficients annulés et les coefficients significatifs. Par contre, elle présente l'inconvénient de modifier les coefficients significatifs.

$$d_{soft}(a,k,l) = \begin{cases} W(a,k,l) - t_+ & \text{si } W(a,k,l) > t_+ \\ 0 & \text{si } t_- \leq W(a,k,l) \leq t_+ \\ W(a,k,l) + t_- & \text{si } W(a,k,l) < t_- \end{cases} \quad (3.11)$$

Une autre modification des coefficients non-significatifs consiste à appliquer une contrainte de "régularisation". Partant du principe que le domaine non-significatif n'est porteur d'aucune information sur le résultat final on cherche parmi toutes les modifications possibles de ces coefficients, la solution qui apporte le minimum d'information. La contrainte de régularisation de Tikhonov introduite par Bobichon & Bijaoui (97) est un processus itératif qui consiste à réduire d'échelle en échelle les variations des plans lissés en les convoluant par un filtre laplacien, puis à recalculer les coefficients non-significatifs associés à ces variations.

La valeur du seuil t_{\pm} dépend d'une part du risque de l'estimateur \tilde{F} choisi et d'autre part du degré de signification que l'on veut accorder au résultat final. Le degré de signification est la probabilité pour laquelle une variation du signal estimé \tilde{F} est détectée comme significative, c'est-à-dire non-issu d'une fluctuation du bruit. Cette probabilité augmente avec la valeur des coefficients positifs et avec l'opposé de la valeur des coefficients négatifs. Un critère de signification donné ϵ permet de fixer la valeur seuil à partir de la densité de probabilité des coefficients associés à la fluctuation du bruit (cf. 3.2.1), $p_{W_B}(x)$:

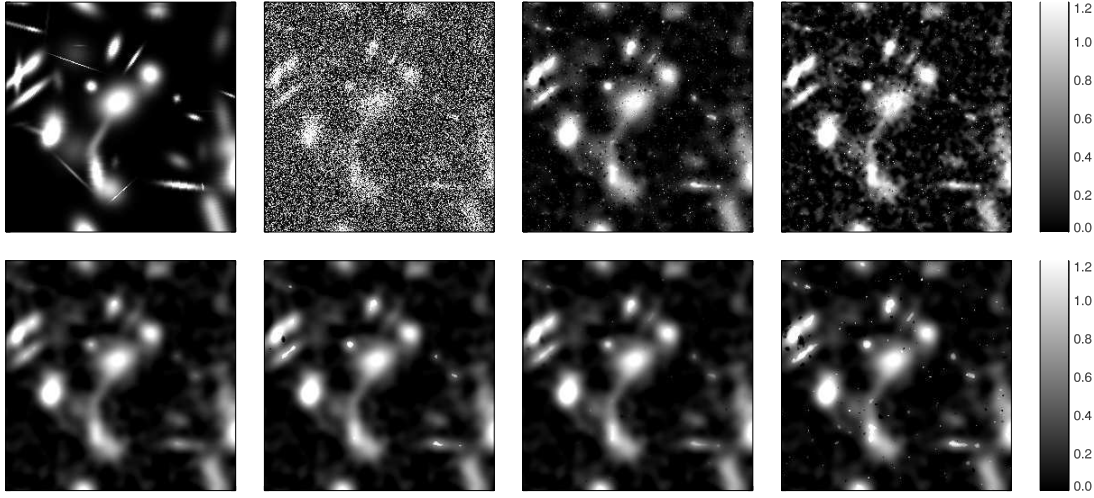


FIG. 3.5 – *Seuillage et filtrage de la transformée en ondelettes adaptés au bruit additif gaussien. Transformée en ondelettes: B-spline, algorithme à trous. (1) Image modèle. (2) Image bruitée (-8dB). (3) Filtrage de Wiener (9.6 dB). (4) Filtrage de Wiener hiérarchique (8.7 dB). (5) Seuillage mou de Donoho (9.8 dB). (6) Seuillage dur de Donoho (10.4 dB). (7) Seuillage mou 3-sigma (10.23 dB). (8) Seuillage dur 3-sigma (8.7 dB).*

$$\begin{aligned}
 t_+ &= \min_{x, x > 0} \left[\int_x^{+\infty} p_{W_B}(u) du \leq \epsilon \right] \\
 t_- &= \max_{x, x < 0} \left[\int_{-\infty}^x p_{W_B}(u) du \leq \epsilon \right]
 \end{aligned} \tag{3.12}$$

Seuillage dans le cas d'un bruit additif gaussien Ce type de seuillage est pratiqué dans l'algorithme de Starck & Bijaoui (1994). Pour un bruit additif gaussien d'écart-type σ , la densité de probabilité des coefficients en ondelettes est elle aussi gaussienne, donc symétrique autour de sa moyenne nulle. On peut alors procéder à un seuillage de la valeur absolue des coefficients avec une valeur seuil de $T = t_{\pm} = k\sigma$.

Mais cet algorithme de seuillage, s'il est robuste, ne minimise pas le risque de l'estimateur \tilde{F} . Donoho & Johnstone (92) ont prouvé qu'il existait dans l'espace des ondelettes, un estimateur de seuillage proche de l'estimateur minimax. Toujours dans le cas d'un bruit additif gaussien, la valeur seuil $T = \sigma\sqrt{2\log N}$ associée à cet estimateur s'exprime aussi en fonction de la déviation standard du bruit, σ , mais aussi du nombre de coefficients indépendants à une échelle donnée N . En pratique, ce seuillage est plus sévère aux hautes résolutions de la transformée en ondelettes, quand un nombre plus élevé de coefficients indépendants augmente la probabilité qu'un seuil fixe soit dépassé par un coefficient associé au bruit.

Sur la figure 3.5 on a comparé les résultats de divers algorithmes de seuillage et de filtrage d'une transformée en ondelettes B-spline calculée selon l'algorithme à trous. Le signal analysé F est une distribution de gaussiennes F , bruitée par un bruit blanc additif gaussien, W . Pour chaque estimation \tilde{F} , on a exprimé le rapport signal-sur bruit final en décibels: $RSB_{dB} = 10 \log \frac{\|F\|}{\|F - \tilde{F}\|}$. On constate que les algorithmes de filtrage de Wiener et de seuillage fournissent des résultats assez similaires en terme de rapport signal-sur-bruit, c'est-à-dire que le flux et les structures globales du signal sont bien restitués dans tous les cas. Les images obtenues par filtrage de Wiener se distinguent par contre des images obtenues par seuillage avec la présence résiduelle d'une texture périodique dans les régions d'émissivité uniforme, ainsi que de petites structures à haute résolution.

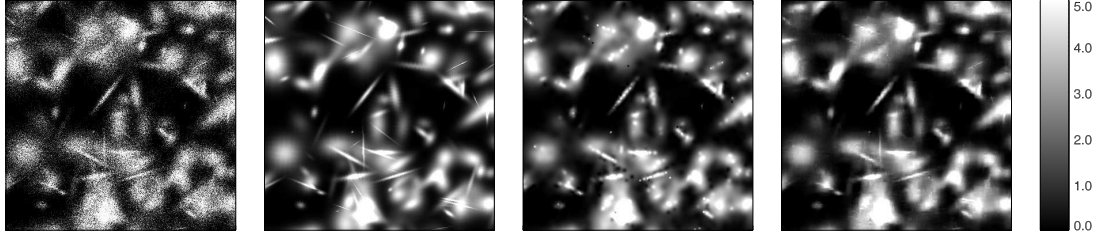


FIG. 3.6 – *Seuillage adapté à un processus de Poisson. (1) Fluctuation de Poisson (5 dB). (2) Image modèle. (3 et 4) Images restaurées par seuillage de la transformée d'ondelettes: (3) algorithme à trous - ondelette B-spline - PDF approchée (18.4 dB). (4) algorithme de Coifman & Donoho - ondelette de Haar - PDF exacte (17.9 dB).*

En effet les algorithmes de filtrage de Wiener dérivent d'une approche bayésienne qui suppose le signal comme étant la réalisation d'un processus gaussien stationnaire. Or les nombreuses structures du signal de la figure 3.5 impliquent des corrélations à différentes échelles qui ne correspondent pas à un processus stationnaire. Parmi les quatre images obtenues par seuillage de la transformée en ondelettes, les deux images obtenues par seuillage dur sont plus structurées à haute résolution. Mais cette structuration correspond souvent à de fausses détections dans le cas du seuillage dur à 3σ . Le seuillage de Donoho fournit un résultat plus satisfaisant en supprimant les fausses détections, c'est l'image obtenue avec un seuillage dur de Donoho qui a le meilleur rapport signal-sur-bruit.

Seuillage dans le cas d'un processus de Poisson. Deux types de seuillages sont envisageables dans le cas d'un processus de Poisson. Soit la densité d'événements n'est pas trop faible ($n > 10$) et on peut utiliser la transformée d'Anscombe pour obtenir un signal bruité par un bruit additif d'écart-type σ . Les algorithmes de seuillage de Starck & Bijaoui et Donoho & Johnstone s'appliquent alors. Soit la densité d'événements peut atteindre de très faibles valeurs, comme c'est parfois le cas en astronomie X, et la valeur seuil doit être calculée à partir d'un critère de signification sur la densité de probabilité des coefficients associés à un processus de Poisson uniforme (cf. equation (3.6)).

L'algorithme de Kolaczyk (1999) repose sur ce principe en attribuant des valeurs seuils dépendant du paramètre de Poisson local aux coefficients en ondelettes. Le calcul de ces valeurs a été effectué en utilisant une approximation des distributions des valeurs de probabilité de l'équation (3.12), pour un processus de Poisson uniforme dans le cas d'une ondelette de Haar. Enfin l'algorithme de Jammal & Bijaoui (1999) propose également un seuillage de la transformée en ondelettes de Haar, mais cette fois avec un calcul exact et analytique de la PDF des coefficients et donc des seuils. C'est finalement cet algorithme que nous avons choisi pour notre problème. Sa description détaillée fera l'objet de la prochaine partie.

La figure 3.6 illustre la restauration d'un signal par deux algorithmes de seuillage adaptés au processus de Poisson. Le premier a été développé par Starck & Bijaoui et utilise des ondelettes B-splines et un algorithme-à-trous. La PDF des coefficients associés au processus stationnaire est approchée. Le second est l'algorithme de Jammal & Bijaoui, utilisant l'ondelette de Haar, l'algorithme de Coifman & Donoho et un calcul analytique de la PDF. Ces deux algorithmes sont plus performants que le lissage adaptatif de la même image (cf. fig 3.1), aussi bien en termes de fidélité des structures au modèle que de rapport signal-sur-bruit global. On constate que le rapport signal-sur-bruit est légèrement supérieur (+0.5 dB) en utilisant le premier algorithme, parce que l'ondelette B-spline est plus adaptée à la forme des structures de l'image que l'ondelette de Haar. Par contre, cet algorithme introduit des artefacts dans la forme des structures quand le paramètre de Poisson est rapidement variable. Par exemple, les filaments superposés à un fond uniforme de l'image (3) sont restitués avec une texture qui ne correspond pas au modèle. Ces artefacts sont dus à l'erreur d'estimation itérative du paramètre de Poisson local, plus grande si la PDF des coefficients associés au processus de Poisson stationnaire n'est pas calculée analytiquement.

3.2.3 Comparaison des algorithmes.

Les caractéristiques générales des divers algorithmes que nous avons passés en revue sont résumées dans les tableaux 3.3 et 3.4. On y a comparé les méthodes d'estimation du bruit, les natures itérative et redondante et le type des algorithmes respectifs. On a aussi mentionné les algorithmes qui permettent d'estimer simultanément une densité locale et un fond à plus grande échelle. Ces derniers sont utiles dans le cas où l'on observe des sources étendues superposées à un fond non uniforme et non modélisable.

Si le nombre d'événements est suffisant pour appliquer la transformée d'Anscombe ($N > 10$), on pourra utiliser les algorithmes d'estimation dans le cas d'un bruit additif (cf. tableau 3.3 et figure 3.5). On a vu que le filtrage de Wiener, de type bayésien, risque alors d'introduire des textures et des artefacts si le processus s'écarte trop de la stationnarité. D'une manière générale, on n'utilisera de méthode bayésienne que si le modèle de signal est adapté. En l'absence d'a priori sur le signal on préférera les méthodes de seuillage de la transformée en ondelettes qui sont robustes et efficaces dans tous les cas.

Dans le cas d'un processus de Poisson impliquant un faible nombre d'événements, on choisira l'un des algorithmes du tableau 3.4. Ne disposant pas d'a priori sur la distribution du paramètre de Poisson on choisira un algorithme fréquentiste. Parmi eux on écartera les algorithmes de lissage et d'échantillonnage adaptatif qui ne procèdent pas à une analyse multiéchelle complète et ne sont pas itératifs. L'intérêt du lissage adaptatif d'Ebeling et al. est la possibilité d'estimer le fond local. Dans notre cas le fond est supposé connu et modélisable. Il sera donc soustrait après le débruitage. Les algorithmes les plus performants et les plus robustes sont encore les algorithmes de seuillage et de filtrage de la transformée en ondelettes. L'algorithme de filtrage de Nowak & Baraniuk procède à un filtrage qui n'est qu'asymptotiquement optimal avec le nombre d'événements, par ailleurs il n'est pas itératif. On lui préférera donc l'un des deux algorithmes de seuillage adaptés au bruit de Poisson. Malgré les artefacts, –effets de bloc–, introduits par l'analyse en ondelettes de Haar, on a finalement choisi l'algorithme de Jammal & Bijaoui, qui procède à une estimation itérative du paramètre de Poisson local en utilisant la PDF exacte des coefficients associés à un processus de Poisson stationnaire.

3.3 Restauration d'image adaptée aux processus de Poisson à faible statistique

Nous détaillons ici un algorithme de restauration d'image adapté aux processus de Poisson à faible statistique, développé par Jammal & Bijaoui (2000) dans le cadre d'une application à l'imagerie gamma en médecine. Dans cet algorithme, la détection des variations du paramètre de Poisson local du signal $I(k,l)$ repose sur le seuillage de sa transformée en ondelettes de Haar. Cette ondelette a été choisie parce qu'elle permet un calcul analytique de la densité de probabilité des coefficients associés à un processus de Poisson stationnaire local. Pour appliquer cet algorithme à l'astronomie X, nous avons choisi d'utiliser une transformée redondante et anisotrope, ainsi qu'une reconstruction régularisée dans le seul domaine non-significatif de la transformée.

3.3.1 Seuillage de la transformée.

Considérons une image $I(k,l)$ résultant d'un processus de Poisson non stationnaire, le principe de cet algorithme est d'isoler les régions de l'espace temps-fréquence pour lesquelles on détecte des variations du paramètre de Poisson avec un niveau de confiance fixé. Ces régions sont isolées en seuillant la transformée en ondelettes de l'image. Connaissant la densité de probabilité des coefficients en ondelettes associés à un processus de Poisson stationnaire local, on considère comme significatif un coefficient en ondelettes dont la probabilité d'être associé à un processus stationnaire est inférieure au niveau de confiance.

La densité de probabilité des coefficients en ondelettes associés à un processus de Poisson stationnaire local, $p_w(w)$ peut s'exprimer par une série d'autoconvolutions de l'histogramme de

l'ondelette, $H_\psi(w)$ (cf. équation (3.6)). Sa transformée de Fourier $\phi_w(\nu)$, aussi appelée fonction caractéristique, s'exprime plus simplement par une série de produits (Jammal & Bijaoui, 1999):

$$\begin{aligned}\phi_w(\nu) &= e^{-\lambda} \left[1 + \sum_{n=1}^{\infty} \widehat{H}_\psi^n(\nu) \frac{\lambda^n}{n!} \right] \\ &= e^{-\lambda} e^{\lambda \widehat{H}_\psi(\nu)} \\ &= e^{\lambda(\widehat{H}_\psi - 1)}.\end{aligned}\tag{3.13}$$

Cette fonction a une expression analytique simple dans le cas d'une ondelette de Haar. En effet, l'histogramme de l'ondelette de Haar s'écrit:

$$H_\psi(w) = \frac{1}{2} [\delta(w-1) + \delta(w+1)],\tag{3.14}$$

et sa transformée de Fourier:

$$\begin{aligned}\widehat{H}_\psi(\nu) &= \frac{1}{2} \int_{-\infty}^{+\infty} [\delta(w-1) + \delta(w+1)] e^{-i2\pi\nu w} dw \\ &= \frac{1}{2} [e^{i2\pi\nu} + e^{-i2\pi\nu}] \\ &= \cos(2\pi\nu).\end{aligned}\tag{3.15}$$

En insérant (3.15) dans (3.13), on en déduit l'expression de la fonction caractéristique des coefficients en ondelettes de Haar associés à un processus de Poisson stationnaire, $\phi_W(\nu) = e^{\lambda(\cos(2\pi\nu)-1)}$ (Jammal & Bijaoui, 1999). Cette fonction étant périodique, on peut la développer en séries de Fourier sur la base $\{e^{-i2\pi\nu n}\}_{n \in \mathbb{Z}}$:

$$\phi_w(\nu) = \sum_{n=-\infty}^{\infty} \langle \phi_w(\nu), e^{-i2\pi\nu n} \rangle e^{-i2\pi\nu n}.\tag{3.16}$$

Le produit scalaire $\langle \phi_w(\nu), e^{-i2\pi\nu n} \rangle$ se calcule sur $[0, \pi]$ pour des raisons de périodicité et de parité:

$$\begin{aligned}\langle \phi_w(\nu), e^{-i2\pi\nu n} \rangle &= \frac{1}{\pi} \int_0^\pi \phi_w(\nu) e^{-i2\pi\nu n} d\nu \\ &= \frac{1}{\pi} \int_0^\pi \phi_w(\nu) \cos(-2\pi\nu n) d\nu \\ &= \frac{1}{\pi} \int_0^\pi e^{\lambda(\cos(2\pi\nu)-1)} \cos(2\pi\nu n) d\nu \\ &= e^{-\lambda} I_n(\lambda),\end{aligned}\tag{3.17}$$

où $I_n(\lambda)$ est la fonction de Bessel modifiée, $I_n(z) = \frac{1}{\pi} \int_0^\pi e^{z \cos \theta} \cos(n\theta) d\theta$. La densité de probabilité des coefficients s'obtient finalement en calculant la transformée de Fourier inverse de $\phi_w(\nu)$ (Jammal & Bijaoui, 1999):

$$\begin{aligned}p_w(w) &= \sum_{n=-\infty}^{\infty} \langle \phi_w(\nu), e^{-i2\pi\nu n} \rangle \delta(w-n) \\ &= \sum_{n=-\infty}^{\infty} e^{-\lambda} I_n(\lambda) \delta(w-n).\end{aligned}\tag{3.18}$$

On remarque qu'il s'agit d'une densité discrétisée qui ne prend que des valeurs entières. Ce résultat était attendu parce que le filtrage d'un processus de Poisson par une transformée en ondelette de Haar est une combinaison de sommes et de différences de valeurs entières. Etant donné un paramètre de Poisson local $\lambda(k,l)$ et un niveau de confiance ϵ , c'est cette densité de probabilité qui permet d'établir un critère de signification pour chaque élément de résolution de la transformée en ondelette du signal $I(k,l)$ (cf. equation (3.12)). Le paramètre de Poisson local $\lambda(k,l)$ étant a priori inconnu et même recherché, le processus de seuillage doit être inséré dans un algorithme itératif et convergent, qui réestime le paramètre de Poisson à chaque étape.

3.3.2 Algorithme de restauration.

L'algorithme de restauration de Jammal & Bijaoui est un algorithme itératif, dans lequel le paramètre de Poisson local $\lambda(k,l)$ est estimé à chaque étape à partir de l'estimateur du signal $\hat{I}(k,l)$. La donnée de ce paramètre et d'un niveau de confiance ϵ permettent de seuiller les coefficients de la transformée en ondelettes de Haar de l'estimateur, $W_{\hat{I}}(s,k,l)$, en lui attribuant une distribution de seuils de signification $t(s,k,l)$ (cf. equation (3.12)). Notons que l'ensemble du processus de restauration porte sur l'estimateur $\hat{I}(k,l)$, ainsi que sa transformée en ondelettes et non sur le signal bruité lui même. L'accord entre les transformées en ondelettes du signal et de l'estimateur est obtenu en contraignant les coefficients associés aux deux transformées, $W_I(s,k,l)$ et $W_{\hat{I}}(s,k,l)$, à rester suffisamment proches de la transformée du signal dans le domaine significatif.

Si l'ondelette de Haar permet le calcul de la densité de probabilité des coefficients, elle est par contre mal adaptée à l'analyse de signaux réguliers parce qu'elle présente de fortes discontinuités et n'a qu'un seul moment nul. La reconstruction de l'estimateur $\hat{I}(k,l)$ à partir d'un seuillage tout ou rien de la transformée ferait intervenir des artefacts caractéristiques de la forme de l'ondelette, en l'espèce des effets de blocs. C'est pourquoi la reconstruction itérative de l'estimateur $\hat{I}(k,l)$ fait intervenir un seuillage régularisé qui modifie les coefficients du domaine non-significatif pour obtenir le signal le plus lisse possible parmi toutes les solutions.

Cet algorithme recherche donc l'estimateur optimal $\hat{I}(k,l)$ de la distribution du paramètre de Poisson $\lambda(k,l)$ compte-tenu, d'une part de l'estimation de $\lambda(k,l)$ dans le domaine significatif et d'autre part, d'une contrainte de régularisation dans le domaine non significatif.

Estimation locale dans le domaine significatif.

La transformée en ondelettes de l'estimateur $\hat{I}(k,l)$ doit reposer sur des estimateurs locaux du paramètre de Poisson $\lambda(k,l)$. Le nombre d'événements observés dans un élément de résolution (k,l) associé au signal bruité est un estimateur non biaisé du paramètre de Poisson local $\lambda(k,l)$. Le choix des éléments de résolution intervenant dans cette estimation est en fait effectué par le seuillage de la transformée en ondelettes du signal bruité $I(k,l)$. Ainsi, on peut procéder à l'estimation locale de $\lambda(k,l)$ en introduisant les coefficients significatifs de la transformée en ondelettes du signal bruité $I(k,l)$ dans la transformée de l'estimateur $\hat{I}(k,l)$. L'accord entre les transformées $W_I(s,k,l)$ et $W_{\hat{I}}(s,k,l)$ dans le domaine significatif est la principale condition à laquelle doit obéir la transformée $W_{\hat{I}}(s,k,l)$ parce que c'est la seule qui la lie au signal bruité observé.

Cette condition s'implémente par un jeu de contraintes non linéaires différenciées entre les coefficients des domaines significatif et non significatif de la transformée de l'estimateur $\hat{I}(k,l)$. Dans le domaine significatif, on force les valeurs des coefficients $W_{\hat{I}}^s(s,k,l)$ à être suffisamment proches des valeurs des coefficients de la transformée du signal $W_I^s(s,k,l)$. Les valeurs possibles sont ainsi délimitées par un intervalle de tolérance, $[W_I^s - \delta t(s,k,l), W_I^s + \delta t(s,k,l)]$, centré sur les valeurs des coefficients $W_I^s(s,k,l)$ et dont la largeur est fonction du seuil de signification local $t(s,k,l)$ et du paramètre $\delta \in [0,1]$. On estime en effet que les valeurs des coefficients peuvent plus ou moins varier en fonction de la fluctuation statistique locale attendue.

$$W_{\hat{I}}^s(s,k,l) = \begin{cases} W_I^s(s,k,l) - \delta t & \forall \{W_{\hat{I}}^s(s,k,l) < -\delta t\}, \\ W_I^s(s,k,l) & \forall \{W_{\hat{I}}^s(s,k,l) \in [W_I^s - \delta t, W_I^s + \delta t]\}, \\ W_I^s(s,k,l) + \delta t & \forall \{W_{\hat{I}}^s(s,k,l) > \delta t\}. \end{cases} \quad (3.19)$$

Dans le domaine non significatif, on force les valeurs des coefficients à être inférieures au seuil de signification local, de sorte qu'aucune information artificielle ne soit introduite par la modification de ces coefficients supposés non porteurs d'information:

$$W_{\hat{\Gamma}}^{ns}(s,k,l) = \begin{cases} -t(s,k,l) & \forall \{W_{\hat{\Gamma}}^{ns}(s,k,l) < -t(s,k,l)\}, \\ W_{\hat{\Gamma}}^{ns}(s,k,l) & \forall \{W_{\hat{\Gamma}}^{ns}(s,k,l) \in [-t(s,k,l), t(s,k,l)]\}, \\ t(s,k,l) & \forall \{W_{\hat{\Gamma}}^{ns}(s,k,l) > t(s,k,l)\}. \end{cases} \quad (3.20)$$

Reconstruction régularisée dans le domaine non significatif.

L'attribution d'une valeur aux coefficients non-significatifs est un problème mal posé qui présente une infinité de solutions. Pour rendre le problème déterministe, on peut introduire une information a priori sur l'état de la solution retenue. La solution choisie pour cet algorithme d'imagerie médicale est celle qui rend l'estimateur $\hat{\Gamma}(k,l)$ et ses approximations successives $F_{\hat{\Gamma}}(s,k,l)$ les plus lisses possibles. Ce choix convient aussi parfaitement à l'imagerie du plasma intra-amas qui est un milieu continu sans discontinuités trop fortes. Il serait par exemple inadapté à l'imagerie planétaire. Chaque approximation doit ainsi subir une contrainte de Tikhonov (1963), qui minimise la norme de son gradient en tout point du plan (k,l) :

$$\frac{\partial}{\partial F_{\hat{\Gamma}}(s,k,l)} \left[\left\| \frac{\partial}{\partial k} F_{\hat{\Gamma}}(s,k,l) \right\|^2 + \left\| \frac{\partial}{\partial l} F_{\hat{\Gamma}}(s,k,l) \right\|^2 \right] = 0. \quad (3.21)$$

On déduit de la dérivation de (3.21) qu'une condition suffisante pour que l'approximation $F_{\hat{\Gamma}}(s,k,l)$ obéisse à la contrainte de Tikhonov est que son laplacien s'annule en tout point du plan (k,l) :

$$\begin{aligned} \frac{\partial}{\partial F_{\hat{\Gamma}}(s,k,l)} \left[\left(\frac{\partial F_{\hat{\Gamma}}(s,k,l)}{\partial k} \right)^2 + \left(\frac{\partial F_{\hat{\Gamma}}(s,k,l)}{\partial l} \right)^2 \right] &= 0, \\ \left[\frac{\partial}{\partial k} \frac{\partial F_{\hat{\Gamma}}(s,k,l)}{\partial k} + \frac{\partial}{\partial l} \frac{\partial F_{\hat{\Gamma}}(s,k,l)}{\partial l} \right] &= 0, \\ L_2 F_{\hat{\Gamma}}(s,k,l) &= 0. \end{aligned} \quad (3.22)$$

L'approximation $F_{\hat{\Gamma}}$ qui vérifie (3.22) peut alors être recherchée en utilisant un algorithme itératif de type Van-Cittert (1931):

$$F_{\hat{\Gamma}}^{(n+1)}(s,k,l) = F_{\hat{\Gamma}}^{(n)}(s,k,l) - \beta L_2(s) F_{\hat{\Gamma}}^{(n)}(s,k,l). \quad (3.23)$$

Enfin pour ramener la contrainte de régularisation dans l'espace des ondelettes, on peut appliquer à (3.23) les opérateurs d'analyse $G_{h,j}$, $G_{v,j}$ et $G_{d,j}$ (cf. (2.49)) associés à un algorithme de filtrage de type multirésolution ou algorithme à trous:

$$\begin{aligned} W_{h, F_{\hat{\Gamma}}^{(n+1)}}(s,k,l) &= W_{F_{\hat{\Gamma}}^{(n)}}(s,k,l) - \beta W_{L_2 F_{\hat{\Gamma}}^{(n)}}(s,k,l) \\ W_{v, F_{\hat{\Gamma}}^{(n+1)}}(s,k,l) &= W_{F_{\hat{\Gamma}}^{(n)}}(s,k,l) - \beta W_{L_2 F_{\hat{\Gamma}}^{(n)}}(s,k,l) \\ W_{d, F_{\hat{\Gamma}}^{(n+1)}}(s,k,l) &= W_{F_{\hat{\Gamma}}^{(n)}}(s,k,l) - \beta W_{L_2 F_{\hat{\Gamma}}^{(n)}}(s,k,l). \end{aligned} \quad (3.24)$$

$$(3.25)$$

La contrainte de Tikhonov sur les approximations se ramène ainsi à un algorithme itératif impliquant les coefficients en ondelettes associés d'une part à ces approximations, $F_{\hat{\Gamma}}(s,k,l)$, d'autre part à la distribution de leurs laplaciens $L_2 F_{\hat{\Gamma}}(s,k,l)$.

La condition de minimisation du gradient (3.21) tend à rendre les approximations $F_{\hat{\Gamma}}(s,k,l)$ et le signal $\hat{\Gamma}(s,k,l)$ uniformes. Elle doit donc être couplée aux contraintes non linéaires (3.19) et

(3.20) pour que le choix de la solution la plus lisse possible s'opère parmi les fonctions qui sont en accord avec la transformée en ondelettes seuillée du signal brut. L'algorithme de Van-Cittert va permettre d'implémenter ce couplage en synthétisant itérativement de nouvelles approximations $\hat{F}_1(s, k, l)$ à partir de coefficients obéissant à la fois à la contrainte de Tikhonov et aux contraintes non linéaires (3.19) et (3.20).

3.3.3 Construction de l'algorithme

Diagramme de construction

Le principe de l'algorithme de Jammal & Bijaoui est de construire itérativement les transformées en ondelettes de l'estimateur $\hat{I}^{(n)}(k, l)$ en tenant compte des contraintes de seuillage et de régularisation des coefficients. Le calcul de la transformée en ondelettes de Haar est implémentable selon l'algorithme multirésolution ou l'algorithme de Coifman & Donoho. La construction de l'algorithme est résumée dans le diagramme 3.8. La première étape est le calcul des coefficients de la transformée en ondelettes du signal bruité $I(k, l)$ ([1] et [2]). Cette décomposition est effectuée de la plus haute à la plus basse résolution par applications successives des opérateurs d'analyse, H_j , $G_{h,j}$, $G_{v,j}$ et $G_{d,j}$ (cf. équation (2.49)). Les étapes suivantes sont les constructions itératives des estimateurs $\hat{I}^{(i)}(k, l)$. Chaque estimateur est construit avec sa transformée en ondelettes, de manière ascendante de la plus basse résolution à la résolution du signal ([3]). A partir du dernier lissé associé à l'estimateur obtenu à l'itération précédente, $\hat{I}^{(i-1)}(k, l)$, les coefficients $W_{\hat{I}(s,k,l)}$ et signaux d'approximation $F_{\hat{I}(k,l)}$ sont obtenus par des reconstructions successives vérifiant l'équation de synthèse des algorithmes de filtrage ([8], cf. 2.50). Pour que cette construction s'effectue en appliquant les contraintes de seuillage et de régularisation, une deuxième itération de type Van Cittert est implémentée à chaque échelle ([10] et [11]) et pour chaque construction d'estimateur. La convergence de ces deux itérations emboîtées est contrôlée par deux critères d'arrêt, $\|W_{\hat{I}(s,k,l)}^{(n)} - W_{\hat{I}(s,k,l)}^{(n+1)}\| < \epsilon_W$ et $\|\hat{I}^{(i)}(k, l) - \hat{I}^{(i+1)}(k, l)\| < \epsilon_I$.

Notons finalement que c'est la construction itérative de l'estimateur $\hat{I}^{(n)}(k, l)$ qui a permis de modéliser à chaque étape la distribution des coefficients associés à un bruit de Poisson stationnaire ([4]) et de définir une distribution de seuils de signification ([5]). La comparaison de ces seuils aux coefficients de la transformée en ondelettes du signal bruité a alors permis de définir des masques délimitant le domaine significatif ([8]) et d'appliquer les contraintes de seuillage et de régularisation. Pour $\epsilon_W = 10^{-3}$ et $\epsilon_I = 10^{-5}$, une vingtaine d'itérations est nécessaire à chaque échelle pour le calcul des coefficients, tandis que l'estimateur $\hat{I}^{(i)}(k, l)$ est généralement obtenu en quatre itérations. L'intérêt des différentes étapes de l'algorithme est illustré par la figure 3.7 où l'on a représenté le signal $\hat{I}(s, k, l)$, reconstruit à partir des transformées en ondelettes $W_{\hat{I}(s,k,l)}$ obtenues successivement après le premier seuillage des coefficients, après la première régularisation et après quatre itérations. Les artefacts apparaissant clairement après le premier seuillage sont atténués par la régularisation. L'image obtenue après une itération présente des structures inexistantes dans l'image de départ parce qu'elle a été obtenue à partir de la transformée en ondelettes seuillée de l'image bruitée. Après quatre itérations, ces structures sont fortement atténuées.

Implémentation de l'algorithme de Coifman & Donoho

Pour appliquer cet algorithme à l'imagerie du plasma intra-amas, nous avons choisi de calculer les coefficients en ondelettes en utilisant l'algorithme de Coifman & Donoho. En effet, la redondance introduite par cet algorithme permet de minimiser le risque de seuillage en moyennant les artefacts de phases distinctes. Elle permet aussi de mieux corrélér la transformée en ondelettes avec le signal en compensant le faible nombre d'échantillons de l'ondelette de Haar. Selon cet algorithme, le calcul de la transformée en ondelettes du signal bruité $I(k, l)$ est effectué par des combinaisons de convolutions horizontales et verticales par les filtres suréchantillonnés \bar{h}_j et \bar{g}_j . Ces filtres sont obtenus en insérant 2^{j-1} zéros entre les deux échantillons des filtres de référence passe-bas, $h = [1, 1]$ et passe-haut, $g = [1, -1]$ associés à l'ondelette de Haar (cf. (2.54)). L'analyse de base à

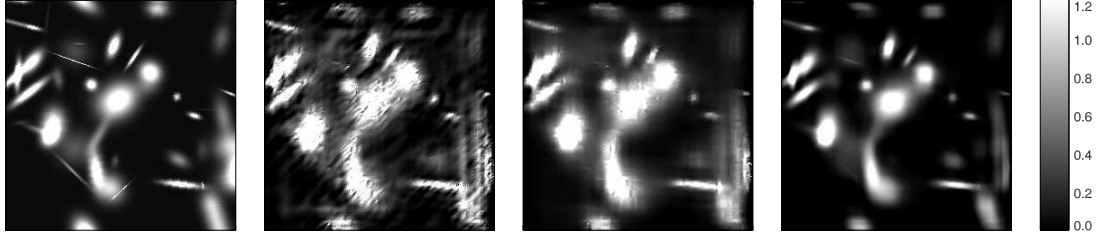


FIG. 3.7 – Débruitage d'une réalisation d'un processus de Poisson par l'algorithme de Jammal & Bijaoui. (1) Paramètre de Poisson. (2) Premier seuillage des coefficients en ondelettes. (3) Première régularisation. (4) Image finale après quatre itérations.

la résolution 2^0 peut se réécrire comme quatre filtrages à deux dimensions associés respectivement aux opérateurs d'analyse H_0 , $G_{h,0}$, $G_{v,0}$ et $G_{d,0}$:

$$\bar{h}_0 \bar{h}_0 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \bar{h}_0 \bar{g}_0 = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, \quad \bar{g}_0 \bar{h}_0 = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \quad \bar{g}_0 \bar{g}_0 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \quad (3.26)$$

Les opérateurs d'analyse H_j , $G_{h,j}$, $G_{v,j}$ et $G_{d,j}$ associés aux résolutions suivantes correspondent à quatre filtrages obtenus en insérant 2^{j-1} zéros aux filtres (3.26).

De même la construction de la transformée en ondelettes de l'estimateur $\hat{I}^{(n)}(k,l)$ fait intervenir les opérateurs de synthèse \tilde{H}_n , $\tilde{G}_{h,n}$, $\tilde{G}_{v,n}$ et $\tilde{G}_{d,n}$, qui correspondent à des convolutions par les combinaisons de filtres suréchantillonnés h_j , g_j , filtres miroirs de \bar{h}_j et \bar{g}_j . La synthèse de base à la résolution 2^0 fait intervenir les opérateurs \bar{H}_0 , $\bar{G}_{h,0}$, $\bar{G}_{v,0}$ et $\bar{G}_{d,0}$, correspondant aux filtrages suivants:

$$\bar{h}_0 \bar{h}_0 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \bar{h}_0 \bar{g}_0 = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \quad \bar{g}_0 \bar{h}_0 = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, \quad \bar{g}_0 \bar{g}_0 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \quad (3.27)$$

Notons enfin que l'opérateur laplacien de l'équation (3.25) correspond lui aussi à un filtrage suréchantillonné et séparable sur les deux axes. Le filtre associé correspond à une double convolution par le filtre gradient, $d = [-1, 1]$ ou encore à une convolution par le filtre laplacien, $l_2 = [-1, 2, -1]$. Le filtrage de base à deux dimensions est de la forme suivante:

$$l_2 l_2(0) = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \quad (3.28)$$

Restriction de la contrainte de régularisation.

Une modification mineure a été apportée à l'algorithme de Jammal & Bijaoui pour cette application astrophysique. Nous avons constaté que l'application de la contrainte de régularisation à l'ensemble des coefficients de la transformée faisait baisser de manière systématique l'amplitude des coefficients significatifs dans la limite des contraintes non-linéaires. En effet, la condition de minimisation du gradient (3.21) tend à rendre les approximations $F_1(s,k,l)$ et le signal $\hat{I}(s,k,l)$ uniformes. Comme les coefficients non-significatifs sont contraints à rester inférieurs au seuil de signification, la discontinuité entre domaine significatif et non significatif est réduite en baissant systématiquement l'amplitude des coefficients significatifs. Nous avons décidé de supprimer cette erreur systématique en limitant la contrainte de régularisation aux seuls coefficients non significatifs. En fait, si cette condition n'avait pas été implémentée par défaut, c'est que la régularisation globale présentait l'avantage de rendre la solution régulière sur tout le plan (k,l) . En imagerie médicale, cet algorithme cherchait à restituer un signal constitué alternativement de creux et de bosses.

La modification d'amplitude des coefficients influait peu sur le flux total de l'image, que l'on ne cherchait par ailleurs pas à restituer de manière optimale. Pour le plasma intra-amas, le signal est principalement constitué d'une source étendue qui constitue une structure positive. La réduction systématique de l'amplitude des coefficients significatifs se traduit par une erreur systématique sur le flux total de l'image, qui est alors sous-évalué. La connaissance du flux de l'image est cette fois primordiale puisque c'est ce paramètre qui permet de quantifier l'émissivité du plasma et d'en déduire ensuite d'autres propriétés comme la masse.

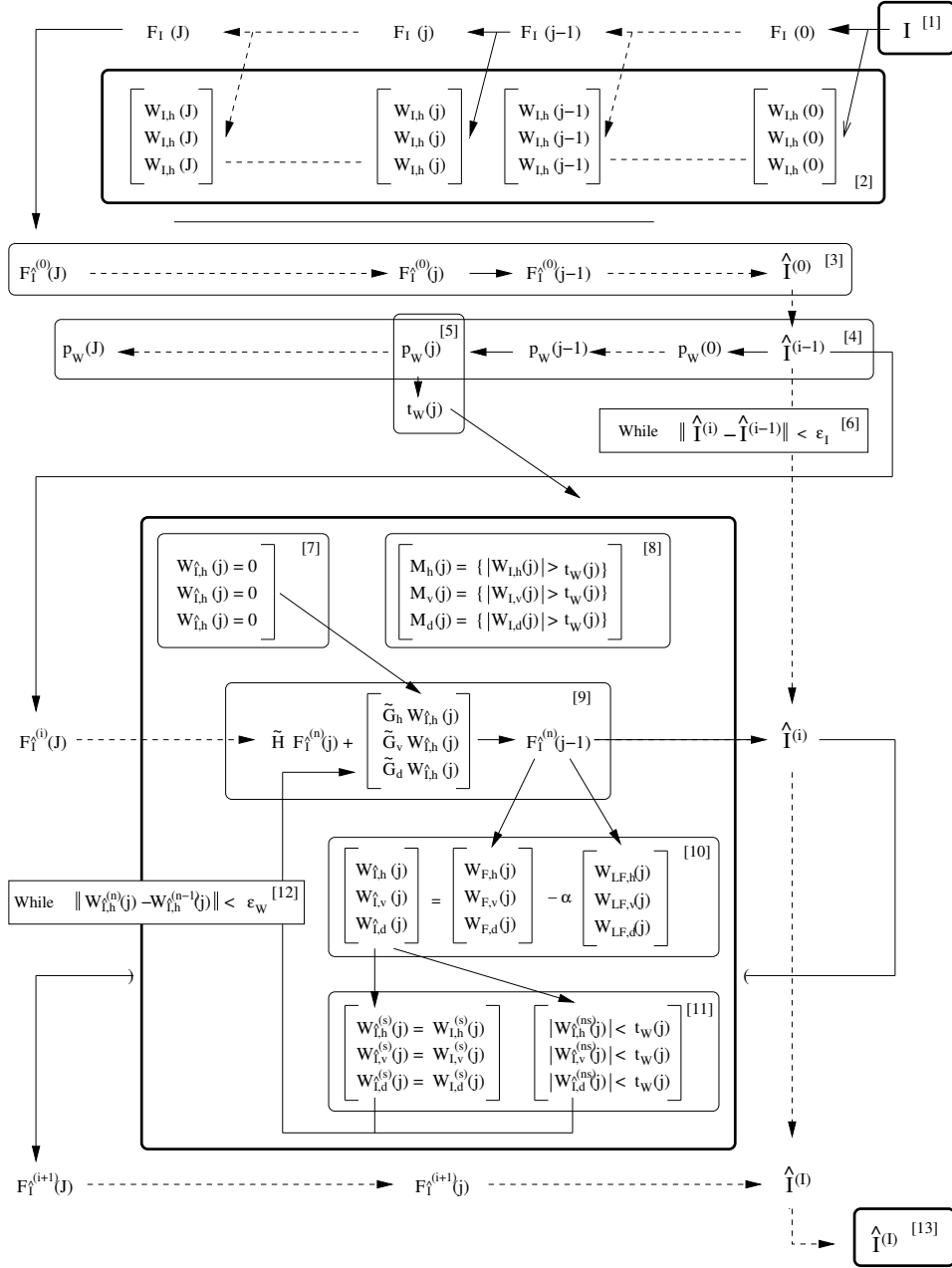


FIG. 3.8 – Algorithme de Jammal & Bijaoui. [1] Image initiale. [2] Calcul de la transformée en ondelettes de l'image initiale. [3] Reconstruction du premier estimateur. [4] Calcul des densités de probabilités des coefficients associés à un processus de Poisson uniforme. [5] Attribution des seuils de signification. [6] Critère de convergence sur le calcul itératif des estimateurs. [7] Initialisation des coefficients. [8] Détermination des masques booléens définissant le domaine significatif. [9] Reconstruction itérative des approximations. [10] Contrainte de régularisation. [11] Contraintes non-linéaires sur les coefficients significatifs et non significatifs. [12] Critère de convergence sur le calcul itératif des coefficients. [13] Estimateur final.

Chapitre 4

Spectro-imagerie X appliquée à l'étude des propriétés du plasma intra-amas.

4.1 Les observables: les paramètres physiques du milieu intra-amas.

Dans le chapitre précédent, nous avons étudié la distribution d'émissivité du plasma intra-amas en considérant la distribution spatiale des événements (k, l, e) comme la réalisation d'un processus de Poisson. L'information spectrale contenue dans ces événements nous permet d'accéder à d'autres paramètres physiques du plasma comme la température, l'entropie, la pression ou la métallicité. C'est pour estimer de manière optimale la distribution spatiale de ces paramètres que nous avons construit un algorithme de spectro-imagerie. Notre algorithme a été développé pour cartographier simultanément les distributions de température et d'émissivité du plasma intra-amas. Nous verrons que les distributions d'entropie et de pression du gaz qui dépendent de ces deux paramètres peuvent facilement en être déduites. La métallicité est par contre un paramètre indépendant de la température et nécessitera des développements ultérieurs.

Pour estimer la distribution d'émissivité, nous avons choisi d'utiliser une approche fréquentiste basée sur le seuillage de la transformée en ondelettes. Rappelons que l'approche fréquentiste a été préférée à l'approche bayésienne parce qu'elle est plus robuste en l'absence d'information préalable sur la distribution de probabilité du signal estimé. En outre, le seuillage de la transformée en ondelettes a permis de construire l'estimateur du paramètre de Poisson en mettant en évidence ses variations spatiales. Le problème se pose de manière assez similaire pour la construction d'un algorithme de spectro-imagerie, où l'on cherche cette fois à estimer la distribution spatiale d'un paramètre θ ou d'un vecteur de paramètres $\boldsymbol{\theta}$. En effet, deux raisons rendent difficile la modélisation a priori de la distribution de probabilité du signal estimé, et conduisent une fois encore à privilégier l'approche fréquentiste. D'une part, la connaissance en ce domaine est très récente, par exemple pour la distribution en température ce sont les satellites X de dernière génération (Chandra, XMM-Newton) qui ont permis de mettre en évidence des structures associées à des phénomènes inattendus auparavant, comme les fronts froids. D'autre part la variété de ces structures, comme les coeurs froids, les fronts froids, les courants de refroidissement ou les ondes de choc, les rend plus difficiles à modéliser a priori que la distribution d'émissivité, qui est toujours continue et assez lisse.

Toutes ces raisons nous ont naturellement conduit à construire un algorithme fréquentiste de spectro-imagerie multiéchelle, utilisant la transformée en ondelettes du paramètre recherché θ ou du vecteur de paramètres $\boldsymbol{\theta}$. Dans cet algorithme, les paramètres $\boldsymbol{\theta}$ sont estimés localement dans différents éléments de résolution de l'espace spatio-fréquentiel, puis leurs variations sont codées par des coefficients en ondelettes. La distribution spatiale optimale des paramètres $\boldsymbol{\theta}(k, l)$ est obtenue à partir de leurs variations spatiales significatives en seuillant leur transformée en ondelettes. Dans une première partie, on étudiera la construction de l'estimateur local $\hat{\boldsymbol{\theta}}$ des paramètres $\boldsymbol{\theta}$, puis dans la partie suivante, on décrira l'algorithme multiéchelle qui permet d'analyser les variations spatiales de ces paramètres.

4.2 L'estimateur local $\hat{\boldsymbol{\theta}}$ et la spectroscopie X.

Cette section est consacrée à la construction d'un estimateur local $\hat{\boldsymbol{\theta}}$ des paramètres $\boldsymbol{\theta}$ du milieu intra-amas en général et en particulier de sa température. Dans une première partie, nous montrerons qu'en constituant une statistique appropriée, T , il est possible de construire un estimateur optimal de variance minimale, puis nous décrirons la méthode du maximum de vraisemblance qui en découle. Nous discuterons ensuite des avantages et inconvénients de cette méthode par rapport à l'estimateur des moindres carrés couramment utilisé en astronomie X. Nous discuterons alors de l'intérêt d'augmenter la résolution spectrale pour l'estimation à travers l'exemple de la technique du rapport de dureté. Finalement nous détaillerons la construction de l'estimateur de température du milieu intra-amas que nous avons implémenté dans l'algorithme de spectro-imagerie présenté dans l'article du chapitre 4.4.

4.2.1 Estimateurs de variance minimum limite (MVB) et borne de Rao-Cramer.

Estimation à un paramètre

Nous avons vu que l'émissivité du plasma intra-amas est modélisable par un modèle d'émission $F(k, l, e)$ (cf. équation (1.2)). Pour construire un estimateur local fréquentiste du paramètre θ , on utilisera un modèle d'émission local $p(e, \theta) \equiv F(e, \theta)$ et on cherchera à lui ajuster une liste d'événements d'énergie e . Pour cela, on pourra associer à chaque événement une pondération $h(e_i)$ et constituer une statistique à partir de l'ensemble des pondérations. La moyenne arithmétique des pondérations constitue par exemple une statistique simple. Nous allons montrer que le choix d'une statistique particulière permet de construire un estimateur optimal, de variance minimale.

Supposons que K événements aient été observés et que l'axe des énergies soit échantillonné en N canaux, la moyenne arithmétique T s'exprime en fonction des I_n événements observés par canal (Bijaoui, 1984):

$$T = \sum_{n=1}^N \sum_{i=1}^{I_n} \frac{h_n(e_i)}{K}. \quad (4.1)$$

Comme l'observation d'un événement e_i dans le canal n de largeur ΔE et d'énergie E_n est la réalisation d'un processus de Bernoulli de probabilité, $p_n(e, \theta) = \int_{E_n - \Delta E/2}^{E_n + \Delta E/2} p(e, \theta) de$, on peut exprimer l'espérance de T par:

$$\begin{aligned} E(T) &= \sum_{n=1}^N E \left[\sum_{i=1}^{I_n} \frac{h_n(e_i)}{K} \right], \\ &= \sum_{n=1}^N I_n \int p_n(e, \theta) h_n(e) \frac{de}{K}. \end{aligned} \quad (4.2)$$

Comme toutes les intégrales associées aux différents canaux sont identiques et compte-tenu de la condition de normalisation, $\sum_{n=1}^N I_n = K$, cette expression se ramène à l'espérance de la fonction h (Bijaoui, 1984):

$$E(T) = g(\theta) = \int p(e, \theta) h(e) de. \quad (4.3)$$

Par ailleurs, la variance de la statistique T s'exprime comme la somme des variances des estimateurs associées aux différents canaux, considérées comme des variables aléatoires indépendantes (Bijaoui, 1984):

$$\begin{aligned} \sigma^2(T) &= \sum_{n=1}^N \sigma^2 \left(\sum_{i=1}^{I_n} \frac{h_n(e_i)}{K} \right), \\ &= \sum_{n=1}^N I_n \left[\int p(e) \left(\frac{h_n(e)}{K} \right)^2 de - \left(\int p(e) \frac{h_n(e)}{K} de \right)^2 \right], \\ &= \frac{1}{K} \left[\int p(e, \theta) h(e)^2 de - \left(\int p(e, \theta) h(e) de \right)^2 \right]. \end{aligned} \quad (4.4)$$

Pour constituer un estimateur $\hat{\theta}$ du paramètre θ , on peut comparer la statistique T et son espérance $E(T) = g(\theta)$, et chercher la valeur de θ telle que $T = g(\theta)$. Ainsi la fonction I inverse de g constitue un estimateur de θ : $\hat{\theta} = I(T)$. Le changement de variable $T = g(\theta)$ permet alors d'approcher au premier ordre la variance du paramètre θ estimé par $\hat{\theta}$ (Bijaoui, 1984):

$$\begin{aligned}\sigma^2(\theta) &\simeq \sigma^2(T) \left[\frac{\partial \theta}{\partial g} \right]^2 \\ &\simeq \frac{1}{K} \frac{\int p h^2 de - \left(\int p h de \right)^2}{\left| \int \frac{\partial p}{\partial \theta} h dx \right|^2}.\end{aligned}\quad (4.5)$$

Nous allons à présent montrer que le choix de la fonction $h(e) = \frac{\partial \log p}{\partial \theta} = \frac{1}{p} \frac{\partial p}{\partial \theta}$ simplifie l'expression (4.5) et minimise la variance approchée de l'estimateur $\hat{\theta}$. Pour cela, calculons tout d'abord pour ce choix, les deux intégrales intervenant dans les termes du numérateur de (4.5):

$$\int p h de = \int \left[\frac{1}{p} \frac{\partial p}{\partial \theta} \right] p de = \frac{\partial}{\partial \theta} \int p de = 0, \quad (4.6)$$

$$\int p h^2 de = \int p \left[\frac{1}{p} \frac{\partial p}{\partial \theta} \right]^2 p de = \int \frac{1}{p} \left[\frac{\partial p}{\partial \theta} \right]^2 de, \quad (4.7)$$

et introduisons-les dans (4.5):

$$\sigma^2(\theta) = \frac{1}{K \int \frac{\left[\frac{\partial p}{\partial \theta} \right]^2}{p} de}. \quad (4.8)$$

Ce choix minimise la variance approchée de l'estimateur $\hat{\theta}$, en effet la fonction $h(e) = \frac{\partial \log p}{\partial \theta}$ annule la dérivée $\frac{\partial \sigma^2}{\partial h}$:

$$\frac{\partial \sigma^2}{\partial h} = \frac{\int -2 \frac{\partial p}{\partial \theta} de}{K \left[\int \frac{\partial p}{\partial \theta} h de \right]^2} = \frac{-\frac{\partial}{\partial \theta} \int p de}{K \left[\int \frac{\partial p}{\partial \theta} h de \right]^2} = 0. \quad (4.9)$$

Cette variance (4.8) peut se réécrire en fonction du nombre d'événements K et de l'information de Fisher F associée à la densité de probabilité $p(e, \theta)$:

$$F = \int \frac{\left[\frac{\partial p}{\partial \theta} \right]^2}{p} de, \quad (4.10)$$

$$\sigma^2(\theta) = \frac{1}{KF}. \quad (4.11)$$

L'estimateur $\hat{\theta} = I[g(\theta)]$ construit en utilisant la statistique T définie selon (4.1) avec la fonction $h(e) = \frac{\partial \log p}{\partial \theta}$ constitue donc l'estimateur du paramètre θ dont la variance est minimale. Cette variance décroît avec le nombre d'événements K disponibles et dépend de l'information de Fisher F apportée par chaque événement pour l'estimation. La construction de cet estimateur repose sur l'inversion de l'équation $T = g(\theta)$, qui n'a de solution unique que si la statistique T est exhaustive, c'est-à-dire construite avec des fonctions h ne dépendant pas explicitement du paramètre θ . Dans le cas contraire, la multiplicité des solutions conduit à estimer le paramètre θ en utilisant un ensemble d'estimateurs de variance minimale, du type $\{\hat{\theta}(\theta_o)\} = \{I[g(\theta, \theta_o)]\}$ dont la variance totale est nécessairement supérieure à celle de l'estimateur unique $\hat{\theta}$ construit avec une statistique exhaustive. La variance (4.11) constitue donc une limite inférieure à la variance des estimateurs du paramètre θ , elle est appelée borne de Rao-Cramer. Un estimateur construit avec une statistique exhaustive du type (4.1) avec $h(e) = \frac{\partial \log p}{\partial \theta}$ atteint cette borne, il est appelé estimateur de variance minimum limite (MVB). Notons enfin que d'après (4.6), les statistiques du type (4.1), qu'elles soient exhaustives ou non, ont une espérance nulle.

Estimation à plusieurs paramètres.

Pour estimer simultanément les p paramètres du vecteur $\boldsymbol{\theta}$, on peut construire un p -estimateur de variance minimale en constituant une p -statistique \mathbf{T} dont chaque composante T_q est une statistique de variance minimale du type (4.1):

$$T_q = \sum_{n=1}^N \sum_{i=1}^{I_n} h_q(e_i), \quad (4.12)$$

avec $h_q(e) = \frac{\partial \log p}{\partial \theta_q}$. L'espérance et la variance de T_q s'expriment par:

$$E(T_q) = g_q(\boldsymbol{\theta}) = \int \left(\frac{\partial \log p}{\partial \theta_q} \right) p(e, \boldsymbol{\theta}) de, \quad (4.13)$$

$$\sigma^2(T_q) = \frac{1}{K} \int \left(\frac{\partial \log p}{\partial \theta_q} \right)^2 p(e, \boldsymbol{\theta}) de. \quad (4.14)$$

L'estimateur MVB du vecteur de paramètres $\boldsymbol{\theta}$ est obtenu en inversant le système de p équations à p inconnues $\{T_q = g_q(\boldsymbol{\theta})\}_{q \in [1, p]}$.

Comme pour l'estimateur à un paramètre, la variance de la statistique T_q s'exprime en fonction des covariances du paramètre θ_q avec les autres paramètres, à partir du changement de variables $T_q = g_q(\boldsymbol{\theta})$ (Bijaoui, 1984):

$$\begin{aligned} \sigma^2(T_q) &= \sum_{m=1}^p \left[\frac{\partial g_q}{\partial \theta_m} \right] \left[\frac{\partial g_q}{\partial \theta_q} \right] \text{cov}[\theta_m, \theta_q] \\ &= \sum_{m=1}^p \left[\int \frac{\partial p}{\partial \theta_m} \left(\frac{\partial \log p}{\partial \theta_q} \right) de \right] \left[\int \left(\frac{\partial \log p}{\partial \theta_q} \right)^2 p de \right] \text{cov}[\theta_m, \theta_q] \\ &= \left[\int \left(\frac{\partial \log p}{\partial \theta_q} \right)^2 p de \right] \sum_{m=1}^p \left[\int \left(\frac{\partial \log p}{\partial \theta_m} \right) \left(\frac{\partial \log p}{\partial \theta_q} \right) p de \right] \text{cov}[\theta_m, \theta_q] \end{aligned} \quad (4.15)$$

De (4.14) et (4.15) on déduit:

$$\sum_{m=1}^p \left[\int \left(\frac{\partial \log p}{\partial \theta_m} \right) \left(\frac{\partial \log p}{\partial \theta_q} \right) p de \right] \text{cov}[\theta_m, \theta_q] = \frac{1}{K}. \quad (4.16)$$

On définit une matrice d'information de Fisher associée à l'estimation des p paramètres par la statistique \mathbf{T} , son terme général s'exprime par:

$$F_{m,n} = \int \left(\frac{\partial \log p}{\partial \theta_m} \right) \left(\frac{\partial \log p}{\partial \theta_q} \right) p de. \quad (4.17)$$

Les p changements de variables $\{T_q = g_q(\boldsymbol{\theta})\}_{q \in [1, p]}$ conduisent à un système de p équations du type (4.16) qui peut se réécrire sous forme matricielle:

$$\mathbf{FV} = \frac{1}{K}, \quad (4.18)$$

où V est la matrice des variances-covariances des paramètres θ_q . Elle s'exprime finalement en inversant la matrice de Fisher:

$$V(\boldsymbol{\theta}) = \frac{1}{K} [\mathbf{F}]^{-1}. \quad (4.19)$$

Enfin la variance associée à l'estimation des paramètres θ_q est déduite des termes diagonaux de V :

$$\sigma^2(\theta_q) = V_{q,q}(\boldsymbol{\theta}). \quad (4.20)$$

4.2.2 Fonction de vraisemblance.

La fonction de vraisemblance, considérée comme une fonction du vecteur de paramètres θ , est maximale pour la réalisation la plus probable de l'échantillon d'événements e_i :

$$L(e_1, \dots, e_n, \theta) = \prod_{n=1}^N \prod_{i=1}^{I_n} p_n(e_i, \theta) \quad (4.21)$$

L'expression (4.12) de la statistique T_q construite avec $h(e) = \frac{\partial \log p}{\partial \theta_q}$ peut se réécrire en introduisant la dérivée logarithmique de la fonction de vraisemblance $L(e_i, \theta)$:

$$\begin{aligned} \frac{\partial \log L}{\partial \theta_p} &= \sum_{n=1}^N \sum_{i=1}^{I_n} \frac{\partial \log p_n(e_i, \theta)}{\partial \theta_q} \\ &= K T_q. \end{aligned} \quad (4.22)$$

Par ailleurs le terme général de la matrice de Fisher $F_{q,m}$ peut-être réécrit en utilisant la fonction de vraisemblance. Pour cela exprimons la tout d'abord en fonction de la dérivée de la fonction h :

$$\begin{aligned} F_{q,m} &= E \left[\left(\frac{\partial \log p}{\partial \theta_q} \right) \left(\frac{\partial \log p}{\partial \theta_m} \right) \right] \\ &= \int \left[\frac{1}{p} \frac{\partial p}{\partial \theta_q} \right] \left[\frac{1}{p} \frac{\partial p}{\partial \theta_m} \right] p de \\ &= \int \frac{\partial p}{\partial \theta_q} \left[\frac{1}{p} \frac{\partial p}{\partial \theta_m} \right] de \\ &= \int \frac{\partial p}{\partial \theta_q} \left[\frac{\partial \log p}{\partial \theta_m} \right] de \\ &= \int \left[\frac{\partial}{\partial \theta_q} \left(\frac{\partial \log p}{\partial \theta_m} p \right) - \left(\frac{\partial^2 \log p}{\partial \theta_q \partial \theta_m} p \right) \right] de \\ &= \int \frac{\partial}{\partial \theta_q} \left[\frac{\partial p}{\partial \theta_m} \right] de - \int \left[\frac{\partial^2 \log p}{\partial \theta_q \partial \theta_m} \right] p de \\ &= \frac{\partial^2}{\partial \theta_q \partial \theta_m} \int p de - \int \left[\frac{\partial^2 \log p}{\partial \theta_q \partial \theta_m} \right] p de \\ &= - \int \left[\frac{\partial^2 \log p}{\partial \theta_q \partial \theta_m} \right] p de \\ &= -E \left(\frac{\partial^2 \log p}{\partial \theta_q \partial \theta_m} \right), \end{aligned} \quad (4.23)$$

.. puis dérivons (4.22) et calculons son espérance:

$$\begin{aligned} E \left(\frac{\partial^2 \log L}{\partial \theta_q \partial \theta_m} \right) &= \sum_{i=1}^{I_n} \sum_{i=1}^{I_n} E \left[\frac{\partial^2 \log p_n(e_i, \theta_q)}{\partial \theta_q \partial \theta_m} \right] \\ &= - \sum_{i=1}^{I_n} \sum_{i=1}^{I_n} F_{q,m} \end{aligned} \quad (4.24)$$

Cette expression traduit l'additivité des informations de Fisher associées à chaque événement du K-échantillon $\{e_i\}$. La fonction de vraisemblance est une fonction du K-échantillon qui permet aussi de définir une matrice d'information de Fisher $F_{K,q,m}$ associée au K-échantillon:

$$F_{K,q,m} = KF_{q,m} = -E \left(\frac{\partial^2 \log L}{\partial \theta_k \partial \theta_m} \right). \quad (4.25)$$

4.2.3 Méthode du maximum de vraisemblance.

La recherche d'un estimateur optimal du vecteur de paramètres $\hat{\theta}$ nous a conduit en 4.2.1, à construire les statistiques T_q associées aux fonctions $h_q(e) = \frac{\partial \log p}{\partial \theta_q}$. Cet estimateur a été obtenu en inversant le système d'équations $T_q = g_q(\theta)$. Puisque l'on a vu que chaque statistique T_q ainsi construite a une espérance nulle, $g_q(\theta) = 0$, le système d'équations $\{T_q = 0\}_{q \in [1,p]}$ constitue lui-même un estimateur implicite du vecteur de paramètres θ . Les statistiques T_q , qui dépendent de la densité de probabilité $p(e, \theta)$ sont alors considérées comme des fonctions des paramètres θ_q et on procède à l'estimation en recherchant les paramètres θ_q qui vérifient $T(\theta_q) = 0$, soit, d'après (4.21):

$$\left\{ \frac{\partial \log L}{\partial \theta_q} \right\}_{q \in [1,n]} = 0. \quad (4.26)$$

Le vecteur de paramètres θ est ainsi obtenu pour une valeur θ_o qui constitue un extrêmu de la fonction de vraisemblance. Si la statistique est exhaustive cette estimation est équivalente à la résolution du système $\{T_q = g_q(\theta)\}$ et la variance du vecteur θ atteint la borne de Rao-Cramer. D'après (4.25) on peut alors réécrire (4.19) en utilisant la matrice $[F_K]$, et exprimer la variance associée à l'estimation des paramètres θ_q en fonction des termes diagonaux de la matrice de variances-covariances V telle que:

$$V(\theta) = [F_K]^{-1} \quad (4.27)$$

Une conséquence de cette série d'égalités est que l'espérance du Hessian de la fonction de vraisemblance est négative en $\theta = \theta_o$ et que l'extrêmu recherché doit être un maximum. La méthode du maximum de vraisemblance consiste à rechercher le vecteur de paramètres θ_o qui maximise $L(\theta)$. Cette méthode fournit un estimateur MVB si la statistique \mathbf{T} est exhaustive et si l'estimateur MVB existe. Si l'estimateur est MVB, la variance des paramètres θ_q est obtenue d'après (4.27) en inversant le Hessian de la fonction de vraisemblance. Dans le cas général où la statistique \mathbf{T} n'est pas exhaustive, on montre que l'estimateur du maximum de vraisemblance est asymptotiquement MVB. En effet si la taille de l'échantillon est suffisante la dépendance en θ_q des statistiques T_q est faible dans l'intervalle de confiance $[\theta_{o,p} - \sigma_q^2, \theta_{o,p} + \sigma_q^2]$ et les statistiques T_q peuvent être considérées comme localement exhaustives.

Il faut alors procéder à des tests numériques pour connaître le comportement asymptotique de l'estimateur et savoir à partir de quelle taille de l'échantillon on peut le considérer comme un estimateur MVB dont les variances σ_q^2 atteignent la borne de Rao-Cramer (4.27).

4.2.4 Estimateur du maximum de vraisemblance et méthode des moindres carrés.

Un estimateur couramment utilisé en spectroscopie X est la méthode des moindres carrés. Cet estimateur est équivalent à l'estimateur du maximum de vraisemblance dans la limite où la statistique est suffisante pour que le spectre d'émission $F(k, l, e)$ soit modélisable par une loi normale dans chaque canal.

Le comptage de photons dans chaque canal n constitue en effet un processus de Bernoulli associé à une variable aléatoire binomiale $B(K, p_i)$. Or le théorème de Moivre-Laplace montre qu'une suite de variables aléatoires binomiales de ce type tend vers la loi de Laplace-Gauss de moyenne $\mu_n = K p_i$ et de variance $\sigma_n = \sqrt{K p_i (1 - p_i)}$. Si cette limite s'applique, la réalisation de la variable aléatoire de Laplace-Gauss, I_n , associée au nombre de photons par canal constitue une observable et la fonction de vraisemblance peut s'écrire en utilisant une nouvelle statistique construite seulement sur un n -échantillon au lieu du K -échantillon considéré jusqu'ici:

$$\begin{aligned}
 \log L &= \sum_{i=1}^N \log \left[\frac{1}{\sigma_n \sqrt{2\pi}} e^{-\frac{[I_n - K p_i(\theta)]^2}{2\sigma_n^2}} \right] \\
 &= -\frac{1}{2} \sum \log [2\pi\sigma_n^2] - \sum_{n=1}^N \frac{[I_n - K p_i(\theta)]^2}{2\sigma_n^2}.
 \end{aligned} \tag{4.28}$$

Dans cette expression, le premier terme est indépendant du vecteur de paramètres θ , la maximisation de $\log L$ revient donc à minimiser le terme:

$$S(\theta) = \sum_{n=1}^N \frac{[I_n - K p_i(\theta)]^2}{2\sigma_n^2}. \tag{4.29}$$

Cette minimisation constitue la méthode des moindres carrés pondérés. On montre que les expressions $S(\theta_k)$ suivent une loi du χ^2 à $(n - p)$ degrés de libertés, p étant la dimension du vecteur de paramètres θ . Cette propriété permet d'associer à l'estimateur des moindres carrés appliqué à la réalisation d'une combinaison de lois normales, un intervalle de confiance indépendant de la loi $p_i(\theta)$. On peut ainsi aisément implémenter un algorithme unique d'estimation des paramètres associées à différentes lois interchangeables du type $F(k, l, e)$. C'est pour cette raison que l'estimateur des moindres carrés est largement utilisé.

4.2.5 Le rapport de durté.

Principe

La technique du rapport de durté est largement utilisée en astronomie X pour construire un estimateur local de température. On définit ainsi deux canaux d'énergie appelés bandes dure et molle qui séparent la liste d'événements (k, l, e) en deux listes, respectivement de haute et de basse énergie. Si K photons sont observés, le nombre de photons observés dans la bande dure suit une loi de probabilité de Bernoulli $B(K, p_H)$, d'espérance $K p_H$ où p_H est la densité de probabilité associée à la bande dure. Le rapport du nombre de photons observés dans la bande dure N_H sur le nombre de photons total K constitue donc un estimateur de la température, il est obtenu avec le "rapport de durté":

$$hr = \frac{N_H - N_S}{N_H + N_S}, \tag{4.30}$$

où N_H est le nombre de photons observés dans la bande dure et N_S le nombre de photons observés dans la bande molle. L'espérance de ce rapport s'exprime en fonction de la densité de probabilité p_H associée à la bande dure:

$$\begin{aligned}
 E(hr) &= \frac{K p_H - (1 - p_H)K}{K p_H + (1 - p_H)K} \\
 &= 2p_H - 1.
 \end{aligned} \tag{4.31}$$

La figure 4.1 est une carte de rapport de durté de l'amas de galaxies A3921 obtenue avec les caméras MOS. Le rapport de durté a été calculé entre les distributions d'émissivité dans les bandes molles (0.3-1.4 keV) et dure (2-12 keV). Les cartes de distributions d'émissivité de chaque bande ont été estimées en appliquant l'algorithme de Jammal & Bijaoui à la distribution poissonienne de photons observée dans chaque bande. Les composantes de fond ont été soustraites de chaque bande avant de calculer le rapport. Dans les régions centrales où l'émissivité de l'amas est importante, les variations de température mises en évidence par le rapport de durté sont proches de celles obtenues en utilisant un véritable algorithme de spectro-imagerie (cf. fig (2.1.1)). En particulier une barre de compression chaude apparaît entre les deux sous-groupes de cet amas bimodal en

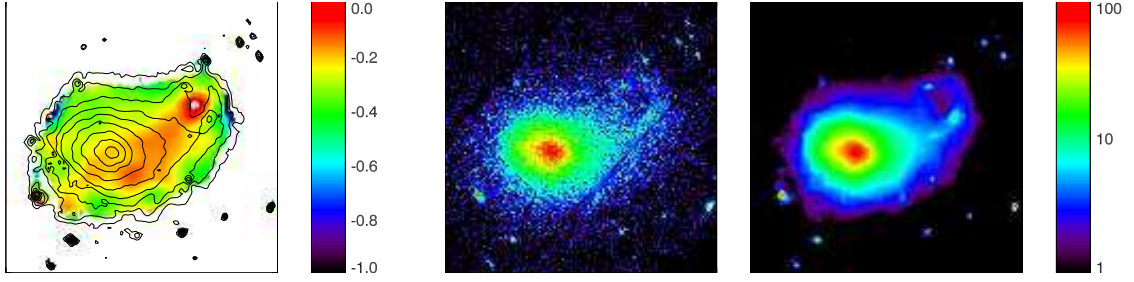


FIG. 4.1 – Image de gauche: Carte de dureté de l'amas de galaxies A3921 obtenue avec les caméras MOS. Les distributions d'émissivité de chaque bande ont été estimées en appliquant l'algorithme de Jammal & Bijaoui à la distribution poissonnienne de photons observée dans chaque bande, puis en soustrayant l'émissivité des composantes de fond. Images de droite: Distribution poissonnienne de photons observée dans la bande molle et carte de distribution d'émissivité de l'amas.

cours de coalescence. Dans les régions externes où l'émissivité de l'amas est faible devant celles des composantes de fond, -région blanche-, l'estimation est biaisée et on constate qualitativement que le résultat obtenu n'a plus de sens physique.

Si cet estimateur est élémentaire à programmer il s'agit d'une méthode empirique dont l'utilisation pratique peut s'avérer complexe et peu efficace. En effet, pour utiliser le rapport de dureté comme un estimateur de température, il faut comparer la statistique observée à un catalogue de densités de probabilités $p_H(\theta)$. Par ailleurs, il ne s'agit pas alors d'un estimateur optimal et il est difficile d'associer un intervalle de confiance à l'estimation. En pratique on utilise le rapport de dureté entre les cartes de densité du plasma construites dans les bandes dure et molle, c'est-à-dire que l'estimation n'est pas optimisée pour la recherche du bon paramètre.

Intérêt de la résolution spectrale.

Le manque de résolution spectrale confère enfin à cet estimateur un défaut pratique supplémentaire pour deux raisons. La première raison est que la définition des intervalles d'énergie définissant les bandes dure et molle doit être optimisée en fonction du modèle d'émission. La figure 4.2 montre un catalogue de spectres d'émission de plasma intra-amas à un redshift de 0.2. On constate que les distributions de probabilités $p_H(\theta)$ associées à la bande dure croissent avec la température, mais qu'inversement les distributions de probabilités $[1 - p_H(\theta)]$ associées à la bande molle décroissent. Intuitivement on comprend que le choix des bandes dure et molle rendra l'estimateur d'autant plus sensible que la dérivée $\frac{\partial p_H(\theta)}{\partial \theta}$ est grande. C'est pourquoi on choisit comme énergie de coupure entre bandes le point d'inflexion où $\frac{\partial p_H(\theta)}{\partial \theta} = 0$. Or ce point dépend du modèle d'émission choisi et de ses paramètres (température, décalage spectral, métallicité, N_H).

La seconde raison est que les différentes composantes du modèle d'émission $F(k, l, e)$ décrit en (1.2) sont forcément imparfaitement modélisées, ce qui entraîne des erreurs systématiques dans la modélisation du flux de photons attendus dans la bande dure. Par exemple le fait de considérer l'émissivité des composantes de fond comme uniforme est une approximation, or une sur- ou sous-estimation d'une de ces composantes conduit à calculer un "rapport de dureté" qui ne vérifie plus l'équation (4.31) et le résultat obtenu est incohérent. C'est ce qui se passe dans les régions externes de la carte de la figure 4.1. Nous montrerons dans la prochaine section que le fait d'augmenter la résolution spectrale de l'estimateur permet de limiter l'influence de ce type d'erreurs systématiques.

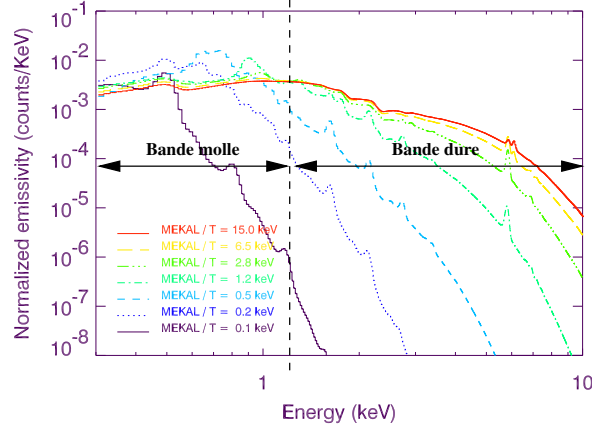


FIG. 4.2 – Définition des bandes dure et molle pour un modèle d’émission de plasma intra-amas à $z=0.2$. On a choisi une coupure de bande à 1.4 keV parce que cette valeur est proche du point d’inflexion $\frac{\partial p_H(\theta)}{\partial \theta} = 0$ pour un amas de 5 keV. Si l’amas était par exemple plus froid on aurait choisi une valeur plus faible.

4.2.6 Application à la spectroscopie X.

Le problème de la résolution spectrale.

La résolution spectrale des spectro-imageurs EPIC varie d’environ 60 eV à 1 keV à 130 eV à 6.4 keV, –raie Fe $K\alpha$ –. Les spectres d’émissions des objets observés peuvent donc être échantillonnées sur quelques centaines de canaux. Cette résolution spectrale est utile pour étudier des propriétés locales du spectre d’émission des sources étudiées comme les raies d’émission des éléments lourds. Mais elle n’est a priori pas nécessaire pour restituer un paramètre θ du milieu intra-amas comme la température, dont la loi de probabilité $p(\theta)$ dépend de manière continue sur l’axe des énergies.

Nous avons discuté au cours de la présentation de la méthode du rapport de dureté l’intérêt d’augmenter la résolution spectrale de l’estimateur pour supprimer le problème de l’optimisation du choix des bandes et limiter l’influence d’erreurs de modélisation. Nous allons à présent quantifier ce dernier point et montrer que nous avons intérêt à utiliser toute la résolution spatiale offerte par les spectro-imageurs EPIC. En construisant une statistique du type (4.1) avec K événements et N canaux contenant chacun I_n événements, on a vu d’après (4.4) et (4.11) que l’on pouvait construire un intervalle de confiance en fonction de la variance du paramètre estimé :

$$\begin{aligned} \sigma^2(\hat{\theta}) &= \sum_{n=1}^N \frac{I_n}{KF} \\ &= \sum_{n=1}^N \sigma_n^2(\hat{\theta}). \end{aligned} \quad (4.32)$$

Ainsi l’influence d’une erreur de modélisation dans un canal sur le résultat final est limitée à la variance du paramètre $\hat{\theta}$, associée à ce canal :

$$\sigma_n^2(\hat{\theta}) = \frac{I_n}{KF} = \frac{I_n}{K} \sigma^2(\hat{\theta}),$$

On a donc intérêt à réduire le nombre de photons I_n reçus par canal “imparfait” en augmentant la résolution spectrale.

Choix de l'estimateur.

Nous allons maintenant chercher à construire un estimateur local de température satisfaisant les contraintes imposées par son implémentation dans un algorithme adaptatif et par les propriétés du milieu intra-amas. L'implémentation de l'estimateur dans un algorithme multiéchelle et la nature fortement variable des processus d'émissivité du milieu intra-amas imposent tout d'abord de construire un estimateur convenant à des statistiques de dynamiques variées. Pour autant cet objectif ne doit pas être atteint au détriment de l'homogénéité de l'estimation puisque des estimations opérées dans différents éléments de résolutions doivent être comparables. Enfin nous venons de discuter de l'intérêt d'augmenter la résolution spectrale de l'estimateur.

Si la méthode du rapport de dureté est une méthode empirique et non optimale, les estimateurs du maximum de vraisemblance et des moindres carrés sont des estimateurs optimaux sous réserve de les utiliser pour des statistiques appropriées. L'estimateur des moindres carrés permet, on l'a vu, une définition des intervalles de confiance à partir de lois du χ^2 tabulées et une implémentation facile de différents modèles d'émission. Mais l'hypothèse d'un modèle d'émission suivant une loi normale dans chaque canal est une hypothèse forte difficilement compatible avec les autres contraintes imposées plus haut. En effet, on montre que l'erreur d'approximation de la loi de Bernoulli $B(K, p_i)$ par une loi de Laplace-Gauss est négligeable pour ($Kp_i > 5$ et $K(1-p_i) > 5$) (Saporta, 90). Cherchant à construire un estimateur homogène et indépendant de la statistique on peut se placer dans le cas courant d'une statistique faible de 500 photons par élément de résolution. La condition $Kp_i > 5$ limite la résolution spectrale à 100 canaux dans le cas le plus favorable d'une égalisation de l'histogramme d'énergie du spectre $F(k, l, e)$, soit: $p_i = 0.01, \forall i$. Cette résolution spectrale est de l'ordre de grandeur de celle des spectro-imageurs EPIC. Malheureusement l'émissivité de la composante thermique du spectre d'émission $F(k, l, e)$ varie de plusieurs ordres de grandeur à haute énergie, et l'égalisation de l'histogramme d'énergie du spectre $F(k, l, e)$ ne peut être valable que sur un faible intervalle de températures. Pour utiliser l'estimateur des moindres carrés il faut donc adapter l'échantillonnage en énergie à l'émissivité et à la température de l'élément de résolution, ce qui complexifie son implémentation dans un algorithme adaptatif et nuit à l'homogénéité de l'estimation. Malgré sa "lourdeur", nous préférons utiliser l'estimateur du maximum de vraisemblance basé sur une statistique de type (4.1), en particulier parce que cet estimateur n'impose pas de contrainte sur l'échantillonnage en énergie et permet d'utiliser la résolution spectrale disponible.

Estimateur de température.

Pour construire un estimateur de température, nous avons décidé de construire une statistique du type (4.1) et d'utiliser la méthode du maximum de vraisemblance pour ses propriétés d'optimalité asymptotiques et son absence de contrainte sur la résolution spectrale. Nous allons maintenant développer les expressions de l'estimateur et de sa variance, mais aussi discuter du nombre de paramètres indépendants posés par le problème.

On rappelle que la probabilité d'émission d'un photon par élément de résolution (k, l, e) est donnée par (1.2) en fonction des densités de probabilités normalisées $S(T, e)$, $B(e)$ et $P(e)$, associées respectivement à l'amas et aux composantes de fond:

$$N_F F(T, e) = N_S S(T, e) + N_B B(e) + N_P P(e), \quad (4.33)$$

où N_F , N_S , N_B et N_P sont les émissivités de chaque composante. Elles vérifient la condition de normalisation suivante:

$$N_F = N_S + N_B + N_P. \quad (4.34)$$

L'émissivité totale N_F étant connue et égale au nombre d'événements, cette condition ramène l'ajustement du modèle d'émission $F(T, e)$ au spectre observé à un problème à trois inconnues, la température du plasma intra-amas dont dépend la composante thermique $S(T, e)$ et les émissivités respectives de deux des trois composantes $S(T, e)$, $B(e)$ et $P(e)$. En première approximation on

peut considérer que l'émissivité des deux composantes de fond $B(e)$ et $P(e)$ est uniforme sur le plan du détecteur et les fixer a priori à partir de l'ajustement préalable d'un modèle d'émission global et isotherme $F(T_o, e)$ associé à l'ensemble des événements d'une observation. Les émissivités N_B et N_P étant connues et fixées, l'ajustement d'un modèle d'émission local $F(T, e)$ est maintenant un problème à une inconnue: la température du plasma intra-amas, T . En pratique, l'hypothèse d'uniformité des composantes de fond n'est qu'une approximation et il serait plus juste de considérer que leurs variations d'émissivités sont faibles. C'est pourquoi nous avons décidé de laisser un degré de liberté supplémentaire au problème en construisant un estimateur à deux paramètres, la température, T et l'émissivité totale, N_F . En laissant libre le paramètre N_F et en supposant constantes les émissivités des composantes de fond on permet un ajustement de l'émissivité de la composante thermique par rapport au fond. Si cet ajustement n'était pas possible l'estimation de température serait biaisée par une mauvaise estimation de l'émissivité des composantes de fond conduisant à ajuster un modèle d'émission $F(T, e)$ de forme erronée. L'émissivité totale N_F étant toutefois proche du nombre d'événements observés nous avons décidé d'introduire un "paramètre d'émissivité" $r = \frac{N}{N_F}$ et d'ajuster le nouveau couple de paramètres $[T, r]$. Ce changement de variables permet de borner l'espace des paramètres et de simplifier la procédure numérique de recherche du maximum de vraisemblance. Les équations (4.33) et (4.34) peuvent se réécrire en fonction des nouveaux paramètres T et r :

$$\begin{aligned} F(r, T, e) &= \left[1 - \frac{r}{N} (N_B + N_P) \right] S(T, e) \\ &\quad + \frac{r}{N} [N_B B(e) + N_P P(e)], \\ &= S(T, e) + \frac{r}{N} [N_B (B(e) - S(T, e)) \\ &\quad + N_P (P(e) - S(T, e))], \end{aligned} \quad (4.35)$$

$$\begin{aligned} N_S &= N_F - N_B - N_P, \\ &= N \left[\frac{1}{r} - \frac{1}{N} (N_B + N_P) \right], \end{aligned} \quad (4.36)$$

et les paramètres r et T qui permettent le meilleur ajustement aux données sont obtenus en maximisant la fonction de vraisemblance à deux paramètres:

$$\log [L(r, T)] = \sum_i \log [F(r, T, e_i)]. \quad (4.37)$$

Enfin, sous réserve que l'estimateur du maximum de vraisemblance se comporte comme un estimateur MVB, la variance des paramètres r et T s'exprime en fonction des termes diagonaux de la matrice des variances-covariances V :

$$V = - \begin{bmatrix} F_1 & \rho \\ \rho & F_2 \end{bmatrix}^{-1} \quad (4.38)$$

avec:

$$F_1 = \frac{\partial^2 \ln(L)}{\partial r^2}, \quad F_2 = \frac{\partial^2 \ln(L)}{\partial T^2}, \quad \rho = \frac{\partial^2 \ln(L)}{\partial T \partial r}.$$

Méthode de recherche du maximum de vraisemblance.

L'algorithme de recherche du maximum de la fonction $L(r, T)$ est détaillé dans l'appendice A de l'article sur la spectro-imagerie X (Chapitre 4.4). Cette recherche devant être implémentée dans un grand nombre d'éléments de résolution, nous avons choisi l'algorithme du gradient conjugué pour sa convergence rapide. Selon cet algorithme, c'est en fait le minimum de la fonction $-L(r, T)$

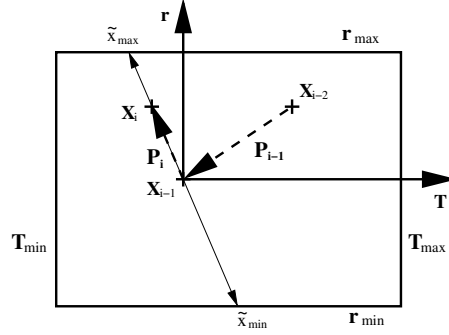


FIG. 4.3 – Recherche du maximum de vraisemblance dans l'espace des paramètres r et T , borné par les valeurs r_{\min} , r_{\max} , T_{\min} et T_{\max} .

qui est recherché. Partant d'un point quelconque de l'espace des paramètres, $\mathbf{X}_o = [r_o, T_o]$, on cherche le minimum de la fonction $-L_1(x)$ qui constitue une coupe de $-L(r, T)$ dans la direction de son gradient en \mathbf{X}_o : $\mathbf{P}_o = \nabla_o L$. On réitère alors le processus dans les directions conjuguées successives \mathbf{P}_i telles que:

$$\mathbf{P}_i^t \mathbf{H}_o \mathbf{P}_{i+1} = 0, \quad (4.39)$$

où \mathbf{H}_o est le Hessian de $L(r, T)$ en \mathbf{X}_o . Enfin un critère d'arrêt stoppe la recherche quand le gradient local est suffisamment faible: $\|\nabla L(r, T)\| < \epsilon$.

Cet algorithme diffère de l'algorithme du gradient dans lequel les directions de recherche successives, \mathbf{P}_i , sont les gradients locaux. Il converge plus rapidement que l'algorithme du gradient parce que les directions successivement explorées sont différentes et non orthogonales entre elles. Nous avons adapté l'algorithme du gradient conjugué à un espace des paramètres borné en limitant les minimisations des fonctions $-L_1(x)$ à des intervalles $[x_{\min}, x_{\max}]$ délimités par le domaine de définition de l'espace des paramètres (cf. figure 4.3). Ce domaine de définition est choisi tel que le maximum de la fonction de vraisemblance soit inclus dans l'espace des paramètres et situé hors de ses bords, de sorte que l'algorithme du gradient conjugué converge.

Si la statistique est bonne, le maximum est ainsi obtenu en quelques itérations. Si la statistique est mauvaise des cas de maxima multiples voire d'absence de maximum peuvent compliquer le problème. Nous avons par exemple constaté des cas où la fonction $F(r, T, e_i)$ présente des maxima multiples à cause de dégénérescences entre la composante thermique $S(T, e)$ et les composantes de fond $B(e)$ et $P(e)$. Dans ces conditions, on observe empiriquement que le "bon" maximum est le plus proche du milieu de l'espace des paramètres, correspondant à la température moyenne de l'amas. Nous avons ajouté une condition qui choisit cette solution le cas échéant. Quand la fonction de vraisemblance ne présente pas de maximum satisfaisant après un nombre d'itérations supérieur au critère d'arrêt de quelques dizaines d'itérations, –pas d'annulation de sa dérivée dans toutes les directions, ou dérivée seconde trop faible–, la valeur "indéfinie" est retournée.

Aspects numériques.

Pour obtenir le minimum des fonctions $-L_1(x)$, on recherche par dichotomie le point x_o qui annule leurs dérivées, $\frac{\partial L_1(x)}{\partial x}$, exprimées en fonction du gradient de $L(r, T)$:

$$\left. \frac{\partial L_1(x)}{\partial x} \right|_{x_o} = \mathbf{P}_i \nabla L(r, T) = 0, \quad (4.40)$$

où les composantes du gradient $\nabla L(r, T)$ s'expriment par les dérivées logarithmiques de (4.37):

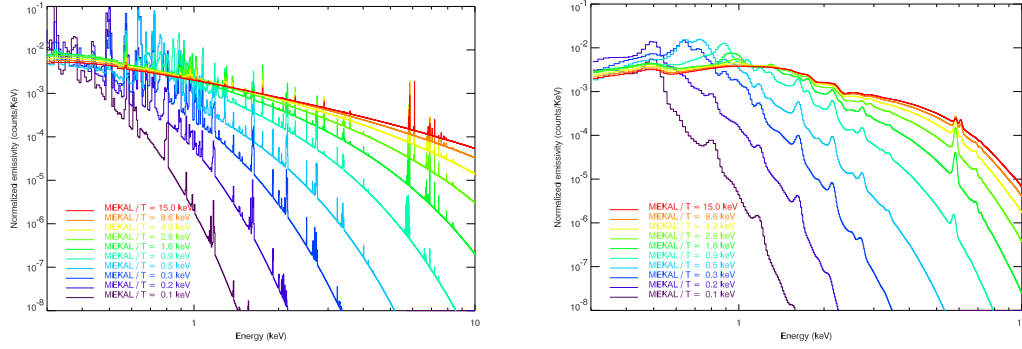


FIG. 4.4 – Bibliothèque de spectres d'émission du plasma intra-amas avant et après convolution par la réponse instrumentale au foyer EPIC.

$$\frac{\partial \log L}{\partial r} = \sum_i \frac{\partial F}{\partial r} \frac{1}{F}, \quad (4.41)$$

$$\frac{\partial \log L}{\partial T} = \sum_i \frac{\partial F}{\partial T} \frac{1}{F}, \quad (4.42)$$

en fonction des dérivées partielles $\frac{\partial F}{\partial r}$ et $\frac{\partial F}{\partial T}$ déduites de (4.35):

$$\frac{\partial F}{\partial T} = \frac{\partial S(T,e)}{\partial T} \left[1 - \frac{r}{N_F} [N_P + N_B] \right], \quad (4.43)$$

$$\frac{\partial F}{\partial r} = \frac{1}{N_F} [N_P [P(e) - S(T,e)] - N_B [B(e) - S(T,e)]]. \quad (4.44)$$

Dans ces deux expressions, les constantes N_F , N_P et N_B , ainsi que les densités de probabilités $P(e)$ et $B(e)$ sont indépendantes des paramètres r et T . Seules les densités de probabilités $S(T,e)$ et $\frac{\partial S(T,e)}{\partial T}$ dépendent de T . Pour pouvoir calculer rapidement les composantes du gradient de $L(r,T)$ en tout point de l'espace des paramètres en fonction de (4.43) et (4.44), nous avons construit une bibliothèque de densités $S_o(T,e)$ et $\frac{\partial S_o(T,e)}{\partial T}$ associées aux composantes thermiques et à leurs dérivées. Les composantes thermiques $S_o(T,e)$ (cf. figure 4.4) sont construites à partir d'un code d'émissivité de plasma optiquement mince (MEKAL, Mewe, 85), puis convoluées par la réponse instrumentale au foyer. Le redshift z , la densité de colonne d'hydrogène N_H et la métallicité Z de l'amas sont fixés. Les composantes $S(T,e)$ et $\frac{\partial S(T,e)}{\partial T}$ sont finalement déduites de $S_o(T,e)$ et $\frac{\partial S_o(T,e)}{\partial T}$ après multiplication par la fonction de vignettage locale $vf(e)$.

Les bibliothèques de densités $S_o(T,e)$ et $\frac{\partial S_o(T,e)}{\partial T}$ sont calculées avec un pas d'échantillonnage de 0.01 keV, très inférieur aux variances attendues sur le paramètre de température. Les spectres $S(T,e)$ intervenant dans le calcul sont obtenus par interpolation B-spline dans la base échantillonnée. Ce sont les bornes de température de la base de spectres qui limitent le domaine de définition de la fonction de vraisemblance. Par exemple, pour un amas de galaxies dont la température moyenne est de 5 keV, l'intervalle des températures exploré s'étend de 0.1 à 30 keV. Quant au paramètre d'émissivité r , il intervient de façon analytique et continue dans l'équation (4.35). Les valeurs de r sont explorées sur l'intervalle $[0,2]$.

Un estimateur de variance minimum limite?

L'estimateur du maximum de vraisemblance présentant asymptotiquement les propriétés d'un estimateur de variance minimum limite (MVB) non biaisé, nous avons testé dans quel régime de

statistique et avec quels modèles d'émission ces propriétés étaient vérifiées. En particulier pouvait-on estimer sa variance par la borne de Rao-Cramer? Un biais est-il observable à faible statistique? Ces tests sont présentés et développés dans l'article du chapitre 4.4.

Le modèle d'émission (4.35) dépendant de trois composantes, $S(T,e)$, $B(e)$ et $P(e)$, nous devons tester les performances de l'estimateur en fonction des émissivités relatives de ces trois composantes. Par ailleurs les composantes d'émission $S(T,e)$ et $B(e)$ associées à l'amas et au fond astrophysique dépendent de la réponse instrumentale locale, -facteur de vignettage-. Nous devons donc aussi tester l'influence du facteur de vignettage sur l'estimation. Pour cela, nous avons simulé des listes d'événements issues d'un modèle d'émission du type (4.35) en tirant au hasard les événements dans un ensemble de distributions de Poisson de paramètre $N_F F(k,l,e)$. Par exemple, pour un modèle d'émission d'amas à 3 keV nous avons simulé des listes d'événements associées à huit valeurs d'émissivités N_F , croissantes de manière logarithmique de 200 à 25600 événements, mais aussi à différentes valeurs d'émissivité des composantes de fond. Pour chaque test, plus de 15000 listes ont été générées.

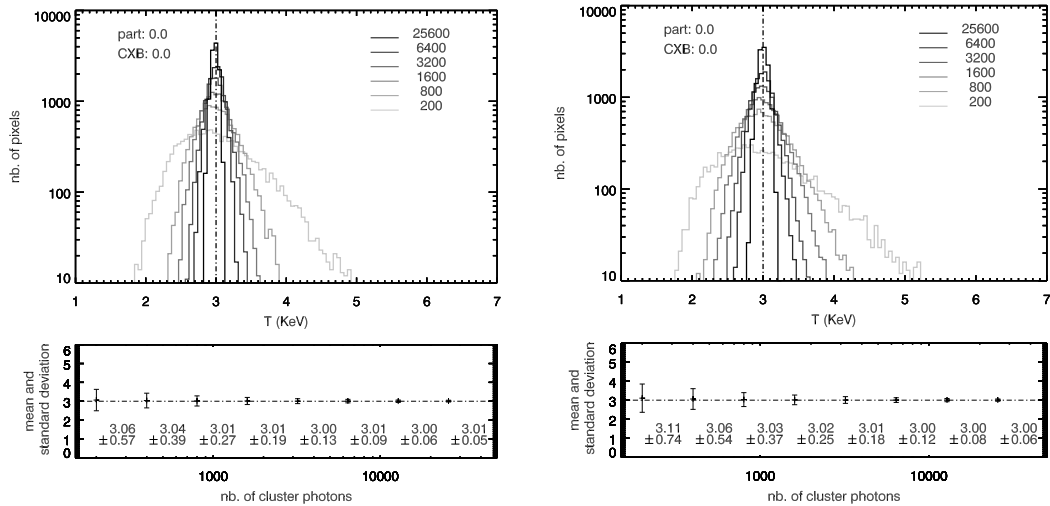


FIG. 4.5 – *Températures estimées à partir d'un modèle d'émission du type $N_S S(k,l,e)$, sans composantes de fond. Gauche: réponse instrumentale uniforme. Droite: réponse instrumentale non uniforme.*

La figure 4.5 présente les distributions des estimations associées à un modèle d'émission sans composantes de fond $N_F F(k,l,e) = N_S S(k,l,e)$, ainsi que leurs moyennes et écarts-types. A gauche, on a supposé la réponse instrumentale uniforme pour les 15000 estimations. A droite on a mélangé des estimations obtenues avec des réponses instrumentales différentes, correspondant aux différentes positions possibles sur le plan du détecteur. Dans tous les cas, les distributions de température sont centrées autour de la valeur de température du plasma, $T = 3$ keV et la valeur moyenne de toutes les estimations obtenues avec plus de 400 photons est de 3 keV à 1 % près. Nous avons par ailleurs vérifié que la variance de ces distributions est très proche de la moyenne des variances MVB calculées à chaque estimation. On peut donc considérer l'estimateur comme un estimateur MVB non-biaisé si le nombre d'événements est supérieur à 400, et ce, que la réponse instrumentale soit uniforme ou non. Ce résultat permet d'utiliser l'estimateur de température en calculant sa variance d'après la borne de Rao-Cramer. Par ailleurs, l'absence de biais corrélé à la réponse instrumentale est crucial pour introduire l'estimateur local dans un algorithme d'imagerie où des estimations calculées dans des éléments de résolution différents doivent être comparables. Si par contre le nombre d'événements est trop faible ($K < 400$), l'erreur d'estimation de la température dépasse 1 % et on constate que l'estimateur est biaisé. Ce biais restant néanmoins de l'ordre de quelques pour-cent et restant faible devant la variance MVB, nous ne

l'avons pas corrigé, préférant limiter l'utilisation de l'estimateur à des listes d'événements dont la statistique dépasse les 500 photons.

Les figures 5 et 6 de l'article du chapitre 4.4 présentent cette fois des estimations associées à un modèle d'émission similaire, mais dont les émissivités des composantes de fond ne sont plus nulles. On constate alors que la variance et le biais de l'estimateur augmentent d'autant plus que l'émissivité relative des composantes de fond par rapport à la composante thermique augmente. Par ailleurs, la variance et le biais de l'estimateur sont plus élevées avec un modèle d'émission comprenant une composante de fond de particules qu'avec une composante de fond astrophysique (fond diffus X et bulle locale, cf. figure 1.4). En fait, l'ajout des composantes de fond a pour effet de réduire la dépendance en température de la bibliothèque de spectres $F(T, k, l, e)$, ainsi que les quantités $\frac{\partial F(T, k, l, e)}{\partial T}$ et les informations de Fisher $\frac{1}{K} \int \left[\frac{\partial F}{\partial T} \right]^2 de$ portées par les événements dans les différents canaux, ce qui augmente finalement la variance de l'estimateur. Cet effet est plus important pour le fond particules que pour le fond astrophysique parce que la densité de probabilité $P(e)$ associée au fond de particules est uniforme spectralement et dépasse de plusieurs ordres de grandeur la densité de probabilité $S(e)$ associée à l'émissivité de l'amas dans sa partie haute énergie (cf. figure 1.4). Nous avons enfin voulu vérifier si, comme dans le cas d'un modèle d'émission sans composantes de fond, la variance des distributions d'estimations était proche de la moyenne des variances MVB estimées pour une statistique dépassant les 500 photons. Là encore il faut distinguer le cas du fond de particules et le cas du fond astrophysique. Les mesures des tableaux 4.6 et 4.7 montrent que la variance de l'estimateur de température est proche de la variance MVB pour le fond astrophysique mais peut la dépasser de 25 % pour un fond de particules d'émissivité double de l'émissivité de l'amas.

Nb. d'événements	400	800	1600	3200	6400	12800	25600
Ecart-type	1.24	0.78	0.50	0.34	0.22	0.16	0.10
Variance MVB	1.14	0.71	0.48	0.33	0.24	0.16	0.12

FIG. 4.6 – *Ecart-type et moyennes des déviations standard MVB pour un modèle d'émission à deux composantes: Plasma intra-amas, $S(e)$ et fond astrophysique par le rayonnement cosmique, $B(e)$. La valeur d'émissivité du fond N_B est double de celle de l'émissivité de l'amas, N_S .*

Nb. d'événements	400	800	1600	3200	6400	12800	25600
Ecart-type	1.67	1.04	0.67	0.45	0.29	0.21	0.14
Variance MVB	1.33	0.79	0.52	0.35	0.25	0.17	0.12

FIG. 4.7 – *Ecart-type et moyennes des déviations standard MVB pour un modèle d'émission à deux composantes: Plasma intra-amas, $S(e)$ et fond de particules, $P(e)$. La valeur d'émissivité du fond N_P est double de celle de l'émissivité de l'amas, N_S .*

Ces tests montrent que nous avons construit un estimateur de température du plasma intra-amas qui, quelle que soit la réponse instrumentale locale, se comporte comme un estimateur MVB pour une statistique supérieure à 500 photons et lorsque les composantes de fond ont une faible amplitude relativement à la composante thermique. Son utilisation dans les cas où les composantes de fond deviennent prédominantes est plus critique parce que l'estimateur ne se comporte alors plus comme un estimateur MVB et parce les dégénérescences de forme des différentes composantes d'émission peuvent le biaiser. Un examen critique des solutions obtenues est alors nécessaire.

Estimateur d'entropie.

L'entropie spécifique du gaz s , -ou entropie par particule-, ou ses variations Δs sont définies à une constante de référence s_o près par (cf. équation (4.8) de la partie I):

$$\begin{aligned}\Delta s &= s - s_o = \frac{3}{2}k \log \left[\left(\frac{T}{T_o} \right) \left(\frac{\rho}{\rho_o} \right)^{\frac{-2}{3}} \right], \\ &= \frac{3}{2}k \log [s'] - s_o,\end{aligned}\tag{4.45}$$

où la quantité $s' = \frac{T}{n^{\frac{2}{3}}}$ est couramment utilisée pour caractériser les variations du paramètre s . En approchant la densité projetée du milieu intra-amas par le carré de son émissivité projetée N_S^2 on peut former un estimateur local \hat{s}' de s' :

$$\hat{s}' = \frac{T}{N_S^{\frac{1}{3}}}\tag{4.46}$$

La maximisation de la fonction de vraisemblance à deux paramètres $L(r, T, e)$ (cf. (4.37)) constituant un estimateur local des paramètres T et N_S nous pouvons simplement en déduire une estimation de s' . Puisque c'est en fait le paramètre d'émissivité normalisé r (cf. 4.36) qui est estimé, l'émissivité du plasma intra-amas s'obtient en fonction de r par:

$$\begin{aligned}N_S &= N_F - N_B - N_P, \\ &= \frac{N}{r} - N_B - N_P.\end{aligned}\tag{4.47}$$

Le paramètre estimé s' est le produit des variables $u = T$ et $v = \left[\frac{N}{r} - N_B - N_P \right]^{\frac{1}{3}}$, de variances respectives:

$$\begin{cases} \sigma_u &= \sigma_T, \\ \sigma_v &= -\frac{1}{3} \left[N_S^{-\frac{1}{3}} \right] \frac{\sigma_r}{r}.\end{cases}\tag{4.48}$$

Sa variance se développe enfin en fonction des variances σ_u et σ_v , –formule de “propagation des erreurs”–, par:

$$\begin{aligned}\sigma^2(s') &= \left[N_S^{-\frac{1}{3}} \sigma_T \right]^2 + \left[-\frac{r}{3} N_S^{-\frac{1}{3}} \frac{\sigma_r}{r} \right]^2, \\ &= N_S^{-\frac{2}{3}} \left[\sigma_T^2 + \left(\frac{T}{3} \frac{\sigma_r}{r} \right)^2 \right],\end{aligned}\tag{4.49}$$

où l'on a négligé le terme de covariance σ_{uv} après avoir vérifié expérimentalement que la matrice de Fisher (4.38) est quasiment diagonale, c'est-à-dire que les estimateurs des paramètres T et N_S constituent des statistiques non corrélées.

En couplant les estimations locales de température et d'émissivité obtenues en maximisant la vraisemblance $L(r, T, e)$, nous construisons simultanément à l'estimateur de la température du plasma un estimateur d'entropie. Ces deux quantités sont complémentaires dans l'étude des propriétés du plasma intra-amas, notamment par ce qu'elles n'ont pas les mêmes échelles de temps de relaxation. L'intérêt de les ajuster simultanément est de ne procéder qu'une seule fois à la maximisation de la fonction $L(r, T, e)$ qui constitue une procédure coûteuse en temps de calcul.

4.3 Algorithme multiéchelle

Ayant construit un estimateur local $\hat{\theta}(k, l)$, nous proposons de restaurer de manière optimale la distribution spatiale du paramètre θ à partir de sa transformée en ondelettes seuillée $W_{\hat{\theta}}(a, k, l)$. La transformation en ondelettes associée à l'analyse multirésolution étant linéaire en échelles, l'idéal

serait d'utiliser un estimateur local $\hat{\theta}(a,k,l)$ unique constitué d'une statistique additive du type (4.1). Ainsi la transformée en ondelettes $W_{\hat{\theta}}(a,k,l)$ pourrait être construite de façon hiérarchique, par exemple de la plus haute à la plus basse résolution en additionnant les statistiques locales appropriées des approximations successives $F_{\hat{\theta}}(a,k,l)$. Malheureusement, l'estimateur $\hat{\theta}_{k,l}(a,k,l)$ dépend lui-même de la position (k,l) sur le plan du détecteur à cause de la réponse instrumentale, –facteur de vignettage–. Nous ne construirons donc pas la transformée en ondelettes de l'estimateur $\hat{\theta}$ mais nous étudierons les variations des estimations θ à toutes les échelles d'une transformée dyadique, en les codant par les coefficients $W_{\theta}(a,k,l)$. Cette démarche est plus lourde parce qu'elle nécessite de procéder à l'estimation dans chaque élément de résolution avant de regrouper en chaque pixel (k,l) les informations obtenues aux différentes échelles. Par ailleurs, puisque les estimations $\theta(a,k,l)$ sont construites avec des estimateurs dépendant de manière non linéaire de l'échelle à cause de la réponse instrumentale, il n'y a a priori pas de relations linéaires entre coefficients $W_{\theta}(a,k,l)$. Pour construire une transformée en ondelettes à partir de ces coefficients il faudra donc introduire une procédure de linéarisation inter-échelles.

4.3.1 Choix de l'algorithme.

Pour étudier les variations spatiales de l'estimation $\theta(a,k,l)$ on pourra se placer dans le cadre de l'analyse multirésolution (cf. section 2.2.1) et définir des cartes d'estimation $\theta(a,k,l)$ à différentes échelles, l'axe des échelles étant échantillonné selon un schéma dyadique. Dans ce cadre, les cartes d'estimation $\theta(a,k,l)$ seront associées à des listes d'événements contenus dans chaque élément de résolution d'échelle a localisé en (k,l) et c'est le choix d'une fonction d'échelle qui déterminera la nature de l'élément de résolution. L'estimateur local $\hat{\theta}(a,k,l)$ étant constitué à partir d'une statistique de comptage, T , on utilisera comme dans l'algorithme d'imagerie adapté au processus de Poisson de Jammal & Bijaoui, une fonction d'échelle indicatrice et on constituera la statistique T à partir des événements contenus dans un carré de largeur a centré en (k,l) . L'ondelette de Haar étant associée à la fonction d'échelle indicatrice, c'est elle que l'on utilisera pour coder les variations spatiales de l'estimation $\theta(a,k,l)$.

Pour minimiser le risque de seuillage de la transformée et compenser la discontinuité de l'ondelette de Haar, on choisira comme pour les algorithmes d'imagerie astronomique d'implémenter un algorithme redondant. Dans le contexte de la spectro-imagerie, la redondance réduira en outre les risques d'estimation. Par ailleurs, pour étudier indépendamment les variations des estimations $\theta(a,k,l)$ aux différentes échelles on devra utiliser un algorithme de filtrage qui n'utilise pas de relations entre coefficients ou approximations d'échelles différentes. Contrairement par exemple à l'algorithme à trous, qui est redondant mais utilise des relations inter-échelles, l'algorithme de Coifman & Donoho satisfait ces deux critères (cf. section 2.3.2). En outre, il est intéressant d'associer cet algorithme à l'ondelette de Haar pour constituer une ondelette anisotrope et analyser les variations des signaux dans trois directions différentes. C'est donc lui que nous utiliserons.

4.3.2 Principe de l'algorithme.

L'analogie entre la restauration optimale du paramètre de Poisson local d'une densité d'événements et d'une estimation quelconque du paramètre $\theta(a,k,l)$ nous a conduit à nous inspirer de l'algorithme de Jammal & Bijaoui pour construire un algorithme de restauration optimale de cartes de température ou d'entropie du milieu intra-amas.

Comme dans l'algorithme de Jammal & Bijaoui, on commence par estimer localement le paramètre θ puis on code les variations de cette estimation en lui associant les coefficients en ondelettes de Haar W_{θ} . La variance de l'estimation à chaque échelle permet d'associer un domaine de signification à ces coefficients. On construit ensuite la transformée en ondelettes $W_{\hat{\theta}}$ de l'estimateur $\hat{\theta}(k,l)$ en minimisant la distance $\|W_{\hat{\theta}} - W_{\theta}\|$ dans le domaine significatif d'une part et en appliquant une contrainte de régularisation aux coefficients du domaine non-significatif d'autre part. Dans l'algorithme de Jammal & Bijaoui, l'intérêt d'utiliser deux transformées distinctes $W_{\hat{\theta}}$ et W_{θ} était de permettre une construction itérative de l'estimateur $W_{\hat{\theta}}$ indépendante de sa réalisation W_{θ} . Ici, l'absence de relation linéaire entre les estimateurs à différentes échelles empêche

de procéder à une estimation itérative. Par contre, l'utilisation des deux transformées distinctes permet de découpler l'ensemble des coefficients W_θ , sans relation linéaire inter-échelles entre eux, de la véritable transformée en ondelettes de l'estimateur final, $W_{\hat{\theta}}$.

Le principe de l'algorithme de construction de cartes en température de dimension $(N1 \times N2)$ est synthétisé par le diagramme 4.8. L'algorithme de construction de cartes d'entropie fonctionne de manière similaire avec l'estimateur d'entropie défini en 4.2.6. On génère tout d'abord des cartes de température à toutes les échelles de l'analyse, $\{T(k, l, j)\}_{j=0, J}$ (1), ainsi que leurs cartes respectives d'erreur, $\sigma_T(k, l, j)$. Pour cela, on procède à des estimations de température dans tous les pixels de largeur $a = 2^j$ centrés en chaque point des cartes $T(j)$. Toutes les cartes d'approximation sont des cartes de même dimension $(N1 \times N2)$ de sorte que l'analyse est quatre fois redondante à chaque échelle. On calcule ensuite les coefficients en ondelettes de Haar horizontaux, $W_{T,h}(k, l, j)$, verticaux, $W_{T,v}(k, l, j)$ et diagonaux, $W_{T,d}(k, l, j)$, associés aux variations spatiales des approximations $T(k, l, j)$ (2). Pour cela, on applique les lissages normalisés en flux $\bar{h}_0\bar{h}_0$, $\bar{h}_0\bar{g}_0$, $\bar{g}_0\bar{h}_0$ et $\bar{g}_0\bar{g}_0$ associés aux opérateurs d'analyse H_j , $G_{h,j}$, $G_{v,j}$ et $G_{d,j}$ de l'algorithme de Coifman & Donoho (cf. section 2.3.2), soit par exemple à la résolution 2^0 :

$$\bar{h}_0\bar{h}_0 = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \bar{h}_0\bar{g}_0 = \frac{1}{2} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, \quad \bar{g}_0\bar{h}_0 = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \quad \bar{g}_0\bar{g}_0 = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \quad (4.50)$$

Les coefficients $W_{T,h,v,d}(j)$ sont ainsi obtenus par une combinaison linéaire de quatre approximations non corrélées:

$$\begin{cases} T_1(j-1) &= T(k, l, j-1), \\ T_2(j-1) &= T(k+2^j, l, j-1), \\ T_3(j-1) &= T(k+2^j, l+2^j, j-1), \\ T_4(j-1) &= T(k, l+2^j, j-1). \end{cases} \quad (4.51)$$

Cette propriété est aussi illustrée par la figure 4.9 dans le cas d'un algorithme de Coifman & Donoho à une dimension, où les coefficients W_j à l'échelle j sont cette fois obtenus par la combinaison linéaire de deux approximations disjointes $F(k)$ et $F(k+2^j)$, à l'échelle $j-1$. On en déduit que la variance du coefficient $W_{T,h,v,d}(j)$ s'exprime comme la somme des variances des quatre approximations $\{T_i(j-1)\}_{i=1,4}$, pondérée par le carré du coefficient $\pm \frac{1}{2}$ associé à chaque approximation:

$$\sigma^2(W(k, l, j)) = \frac{1}{4} \sum_{i=1}^4 \sigma^2(T_i(j-1)). \quad (4.52)$$

Notons que dans le cas des filtrages (4.50) associés à l'ondelette de Haar cette expression conduit à associer la même variance $\sigma^2(T_i(j-1))$ aux trois coefficients $W_{T,h,v,d}(j)$. La donnée des coefficients $W_{h,v,d}(k, l, j)$ et de leur variance $\sigma(W(k, l, j))$ permet de définir un domaine de signification associé à un niveau de confiance $k\sigma(W(k, l, j))$. Ce domaine de signification est délimité par les trois masques booléens, $M_{h,v,d}$, respectivement, horizontal, vertical et diagonal tels que (4):

$$\begin{cases} M_h &= \{|W_h(k, l, j)| > k\sigma_W(k, l, j)\} \\ M_v &= \{|W_v(k, l, j)| > k\sigma_W(k, l, j)\} \\ M_d &= \{|W_d(k, l, j)| > k\sigma_W(k, l, j)\} \end{cases} \quad (4.53)$$

Des approximations de résolution croissante, $F_{\hat{T}}(j)$, sont alors construites à partir du dernier lissé $F_{\hat{T}}(J)$ jusqu'à la carte en température finale \hat{T} , à la résolution 2^0 . Pour cela, on construit itérativement chaque approximation $F_{\hat{T}}(j)$ à partir des coefficients en ondelettes $W_{h,v,d,\hat{T}}(j+1)$ et de l'approximation $F_{\hat{T}}(j+1)$, en appliquant, d'une part, une contrainte de régularisation de Tikhonov aux coefficients non significatifs (6), d'autre part, une contrainte de limitation de la distance entre coefficients significatifs de l'estimateur et coefficients associés aux variations spatiales des approximations $F_T(j)$ (7): $\|W_{h,v,d,\hat{T}}(j+1) - W_{h,v,d,T}(j+1)\| < \delta\sigma_W$, $\delta \in [0, 1]$. L'implémentation de ces deux contraintes est similaire à celle effectuée dans l'algorithme de Jammal & Bijaoui.

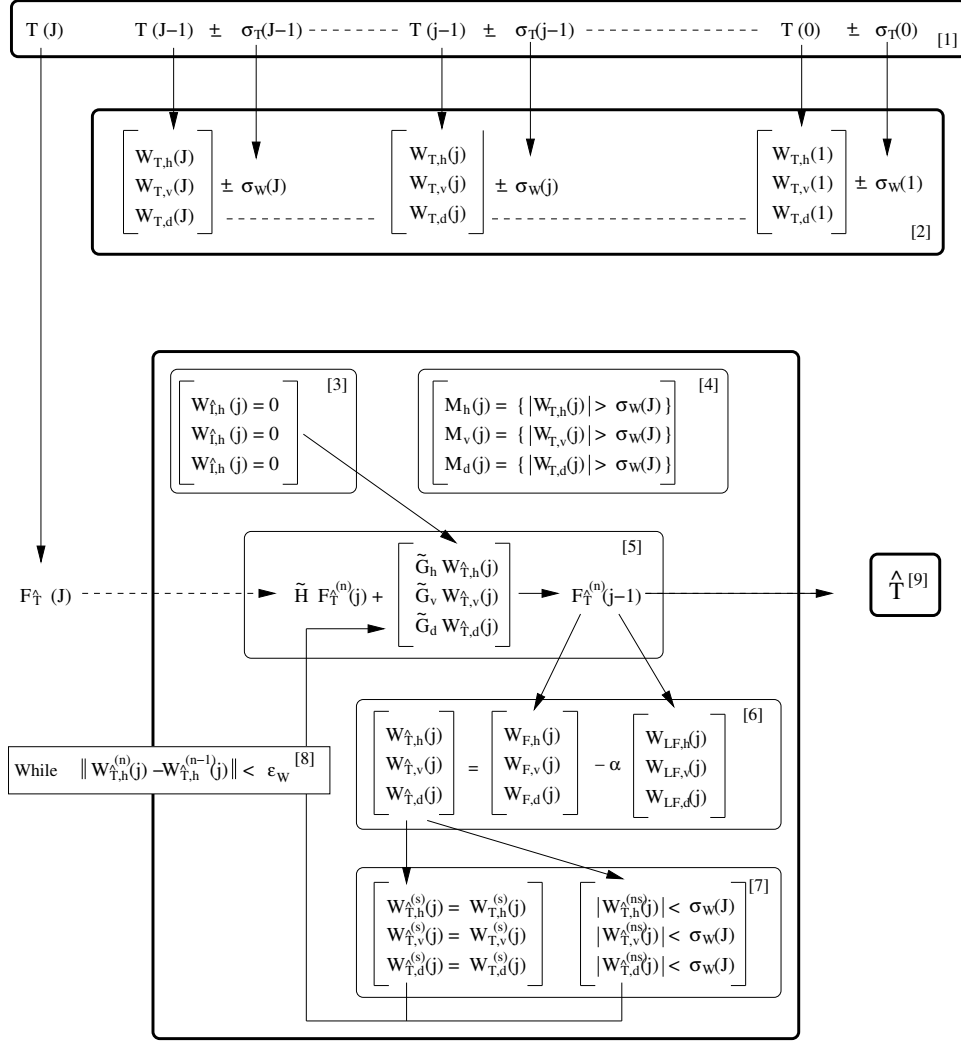


FIG. 4.8 – Algorithme de construction de la transformée en ondelettes des cartes de température. (1) Calcul des cartes de température et cartes d'erreur à toutes les échelles. (2) Calcul des coefficients en ondelettes de Haar à toutes les échelles. (3) Initialisation des coefficients de la transformée de la carte finale. (4) Calcul du masque booléen délimitant le domaine de signification. (5) Reconstruction itérative des approximations. (6) Contrainte de régularisation. (7) Contraintes non-linéaires sur les coefficients significatifs et non-significatifs. (8) Boucle de calcul des coefficients à l'échelle j . (9) Carte en température finale.

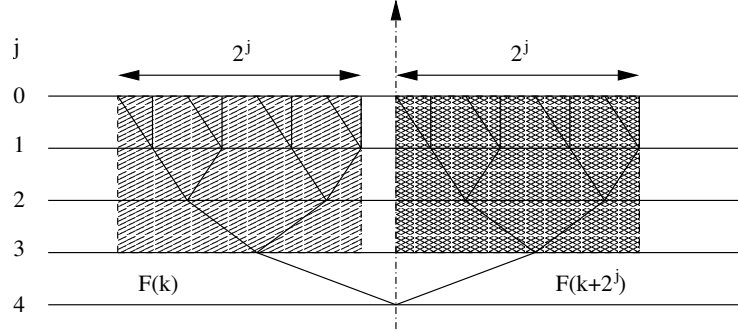


FIG. 4.9 – Superposition du diagramme de fluence associé à l'algorithme de Coiman & Donoho à une dimension et du calcul d'un coefficient W_j à l'échelle j obtenu par une combinaison linéaire des approximations $F(k)$ et $F(k+2^j)$ à l'échelle $j-1$. Le coefficient W_j est obtenu à partir de deux approximations disjointes, non corrélées.

En particulier, la contrainte de Tikhonov (3.22) est appliquée en modifiant itérativement les coefficients non significatifs selon un algorithme de type Van-Cittert (cf. section 3.3.2).

4.3.3 Cartes d'erreur.

On peut grossièrement résumer la construction des cartes en températures $\hat{T}(k,l)$ à la reconstruction de l'image $\hat{T}(k,l)$ à partir de la transformée en ondelettes de l'estimation $T(k,l)$, dont on aurait gardé les coefficients significatifs, mais régularisé les coefficients non-significatifs en leur appliquant une contrainte de Tikhonov. Vue sous cet angle, l'erreur d'estimation commise sur la carte $\hat{T}(k,l)$ est alors majorée par la fluctuation des coefficients significatifs. Pour une statistique uniforme de K événements par pixel, cette fluctuation est déterminée par la borne de Rao-Cramer associée à l'échelle j : $\sigma^2(j) = \frac{1}{K^{4j}F}$. L'écart-type des coefficients est donc divisée par 2 à chaque échelle. C'est pourquoi la fluctuation des coefficients significatifs associés à l'estimation locale $\hat{T}(k,l)$ est essentiellement due aux coefficients significatifs de plus haute résolution associés au pixel (k,l) . On peut donc majorer l'erreur d'estimation commise dans un pixel (k_o, l_o) de la carte $\hat{T}(k,l)$, en additionnant les variances $\sigma_W(j, k_o, l_o)$ associées aux coefficients $W_{h,v,d}(j, k_o, l_o)$, j étant la plus haute résolution du domaine significatif:

$$\sigma_{\hat{T}}(k_o, l_o) = 3 * \sigma_W(j, k_o, l_o), \quad (4.54)$$

La carte d'erreur $\sigma_{\hat{T}}$ ainsi obtenue est indicative des fluctuations d'erreurs associées aux estimations à différentes échelles et permet par exemple une comparaison avec les résultats d'algorithmes adaptatifs dans l'espace direct. Pour autant elle doit être interprétée avec précaution parce que l'information sur la taille de l'élément de résolution qui a conduit à l'estimation locale n'est pas fournie. Par exemple une estimation locale $\hat{T}(k,l) = T_o \pm \sigma_{\hat{T}_o}$ peut être obtenue par des coefficients à grande échelle et ne fournit pas d'intervalle de confiance ni même d'estimation de température dans l'élément de résolution (k,l) à l'échelle 0.

4.3.4 Implémentation de l'algorithme de spectroscopie.

Calcul des approximations.

La procédure de calcul des cartes en température $T(k,l,j)$ à l'échelle j est schématisée par le diagramme 4.10. La liste des événements (k,l,e) a préalablement été rangée dans un tableau échantillonné en position $(N_1 \times N_2)$ et énergie (N_e) . Puisque le plan focal EPIC est composé de trois caméras, on dispose de trois listes d'événements distinctes, et on constitue en fait trois cubes d'événements juxtaposés (1). Un échantillonnage de $128 \times 128 \times 100$ est généralement un bon

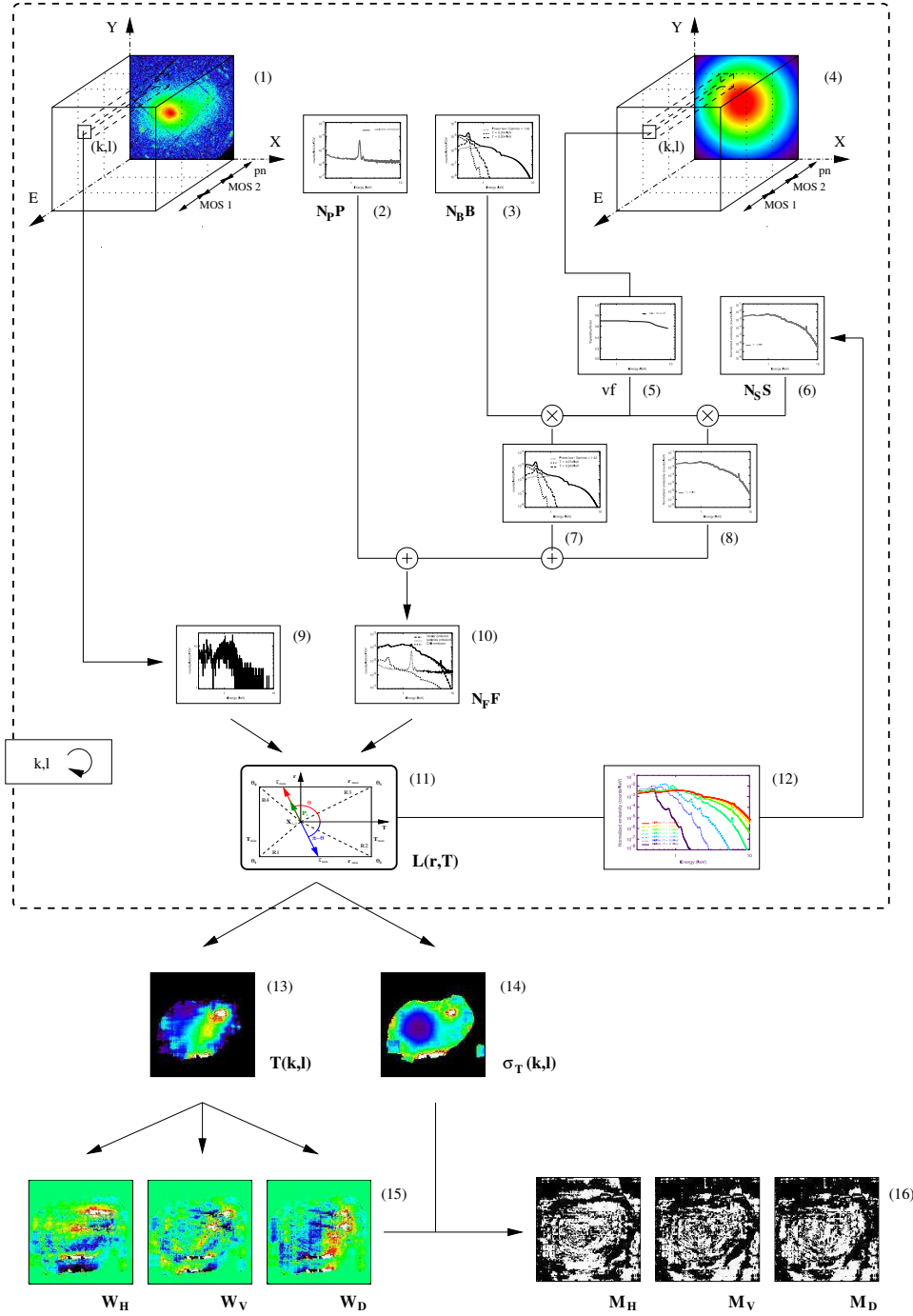


FIG. 4.10 – Calcul des coefficients en ondelettes à l'échelle a . (1) Cube d'événements échantillonné en position (k,l) et énergie (e) . (2) et (3) Spectre d'émission particules et fond astrophysique. (4) Cube de réponses instrumentales, –facteur de “vignettage”–. (5) facteur de vignettage à la position (k,l) . (6) Spectre d'émission de l'amas à la température T . (7) et (8) Spectres corrigés du facteur de vignettage. (9) Liste d'événements à la position (k,l) . (10) Modèle d'émission à la température T . (11) Algorithme de recherche du maximum de la fonction de vraisemblance $L(r,T)$. (12) Bibliothèque de modèles d'émission du plasma intra-amas en fonction de la température. (13) et (14) Carte de température et barres d'erreur à l'échelle a . (15) et (16) Coefficients en ondelettes horizontaux, verticaux et diagonaux et masques de signification associés.

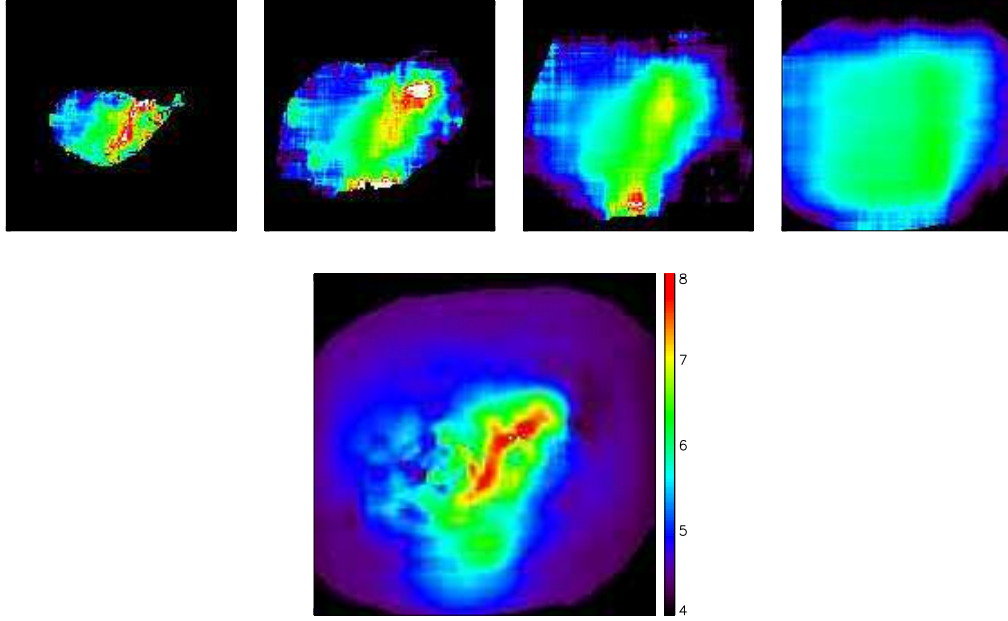


FIG. 4.11 – Images du haut, de gauche à droite: Cartes de température de A3921 sur 4 échelles obtenues avec les caméras MOS. Image du bas: Carte de température finale (échelle de températures en keV).

compromis entre résolution spatiale et statistique pour des sources étendues avec une statistique de 10^5 événements. Cet échantillonnage correspond à un pixel spatial de 14 secondes d'arc. En outre, on constitue un autre tableau associé à la réponse instrumentale locale, -facteur de vignettage-. Ce tableau est échantillonné comme celui des événements (4).

Pour calculer la valeur de température associée au pixel (k,l) de la carte $T(k,l,j)$ à l'échelle j , on définit un élément de résolution de largeur $a = 2^j$, centré en (k,l) , et on extrait, d'une part la liste échantillonnée en énergie d'événements contenue dans cet élément de résolution (9), d'autre part la réponse instrumentale locale moyenne pondérée par la position de chaque événement de la liste extraite (5). La réponse instrumentale locale étant modélisée, on peut générer un modèle d'émission $F(k,l,e)$ du type (1.2) en additionnant la composante thermique $S(e)$ (8) et les composantes de fond astrophysique $B(e)$ (3) et de particules $P(e)$ (2). Les spectres $S(e)$ et $B(e)$ sont obtenues en multipliant les spectres au foyer $S_o(e)$ et $B_o(e)$ (3 et 6) par la réponse instrumentale locale (5). Le modèle d'émission $F(k,l,e)$ (10) peut alors être itérativement ajusté au modèle en maximisant la fonction de vraisemblance à deux paramètres $L(r,T,e)$ (11) (cf. équation (4.37)). Pour cela, on modifie la composante thermique $S(e)$ associée au paramètre de température, en la sélectionnant dans une bibliothèque de spectres (12) préalablement construite (cf. section 4.2.6). La fonction de vraisemblance $L(r,T,e)$ étant maximisée pour les paramètres r_o et T_o , on calcule les variances σ_r et σ_T des deux paramètres en inversant la matrice de Fisher locale (cf. 4.38). La carte $T(k,l,j)$ (13) et la carte d'erreur $\sigma_T(k,l,j)$ (14) sont finalement obtenues en répétant le processus sur tous les pixels (k,l) . Elles permettent de calculer les coefficients $W_{h,v,d,T}(k,l,j)$ (15), de leur appliquer un seuil de signification, puis de définir les limites du domaine significatif (16).

Application à l'amas de galaxies A3921.

La figure 4.11 illustre la reconstruction multi-échelles d'une carte de température du plasma intra-amas pour l'amas de galaxies A3921. Cet amas, constitué de deux sous-groupes en interaction (cf. figure 4.1), a été observé par le satellite XMM dans le cadre d'un programme d'observation d'amas de galaxies en cours de coalescence. On a représenté les approximations $\{T(k,l,j)\}_{j=0,J}$

obtenues avec les caméras MOS sur 4 échelles et la carte finale reconstruite à partir du domaine significatif de la transformée en ondelettes $W_{\hat{\Lambda}}$. Pour obtenir les cartes d'estimations $\{T(k,l,j)\}_{j=0,J}$, on a ajusté un modèle d'émission $F(k,l,e)$ du type (1.2) avec une composante thermique $S(T,e)$ et deux composantes de fond $P(e)$ et $B(e)$. La bibliothèque de spectres de la composante thermique $S(T,e)$ est constituée en fixant le décalage vers le rouge, la densité de colonne d'hydrogène et la métallicité à $z = 0.096$, $N_H = 2.94 \cdot 10^{20} \text{ cm}^{-2}$ et $Z = 0.3Z_{\odot}$, respectivement. Les émissivités N_P et N_B des composantes de fond sont fixées en ajustant le modèle $F(k,l,e)$ à l'ensemble des événements, la composante thermique associée à la température moyenne de l'amas étant supposée uniforme. La statistique minimale pour procéder à l'estimation de température a été fixée à 1000 événements par élément de résolution. La distribution d'émissivité de l'amas décroissant continûment du centre de chaque sous-groupe vers les bords du champ de l'image, cette limite fixe à chaque résolution un contour au-delà duquel les températures ne sont plus estimées, –parties noires des images–. La carte à grande échelle permet de fixer la température moyenne de l'amas à $5.1 \pm 0.6 \text{ keV}$. Les cartes à plus haute résolution laissent apparaître de plus en plus de détails au fur et à mesure que la résolution augmente. En particulier, on observe à toutes les échelles une élévation de température entre les deux sous-groupes, en accord avec l'estimation du rapport de dureté de la figure (4.1). Les températures de cette région intermédiaire, qui correspond à une zone de choc, –cf. section 2.1.1–, peuvent en effet atteindre voire dépasser 7 à 8 keV. Enfin, la carte de plus haute résolution montre que la structuration en température du coeur du sous-groupe central est assez perturbée. Les cinq estimations $\{T(k,l,j)\}_{j=0,5}$ contiennent donc des informations de nature distincte que la carte finale regroupe au mieux en fonction de la statistique disponible.

Si la figure 4.11 montre l'intérêt d'utiliser un algorithme adaptatif pour regrouper des informations à différentes échelles, elle montre aussi qu'une grande partie de l'espace spatio-fréquentiel est inutilisée par cet algorithme du fait de la statistique insuffisante à haute résolution. Ce défaut est dû à l'inadaptation de l'échantillonnage des estimations à la distribution d'émissivité des amas qui décroît fortement du centre au bord. Nous avons choisi d'échantillonner le plan du détecteur par des grilles uniformes parce que ce choix est le plus simple et n'implique pas d'a priori. Néanmoins, on peut s'interroger sur l'opportunité d'adapter l'échantillonnage du plan du détecteur aux variations d'émissivité des amas, qui peuvent atteindre plusieurs ordres de grandeur sur quelques minutes d'arc.

4.3.5 Simulations d'observations.

Le comportement de l'algorithme de spectro-imagerie a été testé sur des observations simulées de deux types. Pour quantifier les erreurs d'estimation commises par l'approche multi-échelles, nous avons simulé l'émission d'un plasma intra-amas à deux dimensions localement isotherme et étudié le comportement de l'algorithme en fonction de la statistique. De manière plus qualitative nous avons par ailleurs simulé l'observation d'un amas de galaxies en cours de coalescence dans des conditions proches des observations réelles, avec un modèle d'émission multi-composantes dans lequel les paramètres caractéristiques du plasma sont projetés dans le plan du ciel à partir d'une simulation numérique à trois dimensions. Ces simulations sont détaillées dans l'article présenté dans la section 4.4.

Cas d'un modèle d'émission isotherme à deux dimensions.

La figure 7 de l'article de la section 4.4 illustre des cartes en températures obtenues à partir d'observations simulées d'un plasma intra-amas isotherme, composé de quatre régions de températures distinctes. Le modèle d'émission est un modèle isotherme de type $F(k,l,e)$ avec deux composantes de fond d'émissivité égale à 5 photons par pixel, et une composante thermique localement variable dont on fixe la température a priori à partir d'une distribution de température $T(k,l)$ quelconque. Les observations sont simulées en tirant au hasard dans chaque élément de résolution (k,l,e) une liste d'événements constituant la réalisation d'un processus de Poisson dont le paramètre est donné par le modèle d'émission $F(k,l,e)$. La distribution d'émissivité de l'amas est une gaussienne centrée au centre du champ, proche d'une distribution de type modèle β . La distribution de température

du plasma est affichée sur l'image du haut à gauche. Elle comprend un disque central à 5 keV, dans lequel on a pratiqué deux excisions pour placer deux régions elliptiques, respectivement à 4 et 6 keV. La température de la région à l'extérieur du disque central est de 3 keV. Cette géométrie en émissivité et température se rapproche de celle d'un amas avec un fort gradient de température centre-bord, -5 keV dans sa région centrale, 3 keV au bord-, comprenant par exemple une région de choc chaude et un coeur froid dans sa partie centrale. Trois observations ont été simulées. Elles correspondent à des émissivités totales de 10^5 , 10^6 et 10^7 photons. On constate qualitativement que les cartes sont restaurées avec d'autant plus de détails que la statistique augmente: avec 10^5 photons seule la structure circulaire à grande échelle de l'amas apparaît, alors qu'avec 10^7 photons les quatre régions de températures différentes sont bien distinguées avec des bords francs correspondant à des structures à petite échelle.

Pour étudier plus en détail le comportement de l'algorithme nous avons affiché sur la figure 8 les profils de température correspondant aux coupes en pointillés sur les cartes de la figure 7 ainsi que les erreurs d'estimation associées. On constate bien qu'en augmentant la statistique on résoud de manière significative les bords francs associés aux frontières du disque et des régions elliptiques centrales. En effet, l'amplitude des barres d'erreur devient moins élevée que celles des variations de températures de ces frontières. Mais on constate aussi que le profil obtenu avec la statistique maximale de 10^7 photons n'est pas en accord avec le modèle d'émission, qui sort des barres d'erreur au voisinage des bords francs. Pour interpréter ce désaccord, on a superposé aux profils de température restaurés de la figure 8, les profils issus des images associées aux deux reconstructions de la transformée en ondelettes de l'image ayant servi de modèle de distribution de température. La première reconstruction a été obtenue après un seuillage identique à celui appliqué dans l'algorithme de spectro-imagerie. Elle se rapproche du modèle initial. La seconde a aussi été obtenue avec le seuillage pratiqué dans l'algorithme de spectro-imagerie, mais après avoir pondéré les températures associées aux différents éléments de résolution des approximations par l'émissivité du plasma. Le profil de cette seconde solution se rapproche du profil obtenu par l'algorithme de spectro-imagerie et est de fait inclus dans les barres d'erreur. La première information livrée par ces comparaisons de profils est que le désaccord constaté entre le profil restauré par l'algorithme de spectro-imagerie et le modèle est dû à l'algorithme multi-échelle et non à un défaut de l'algorithme de spectroscopie.

En fait, le profil restauré est en accord avec le modèle seuillé et restauré avec une pondération des températures, parce que l'algorithme de spectro-imagerie effectue une telle pondération en attribuant à chaque élément de résolution des cartes $T(k,l,a)$ une température proche de celle du plasma dans les régions de plus forte émissivité de l'élément de résolution. La surestimation des températures du profil au voisinage des bords francs s'explique ici de la manière suivante. D'une part la température des différents éléments de résolution est surestimée à cause du profil d'émissivité qui favorise un étalement de la température du centre de l'amas (5 keV) par rapport à celle du bord (3 keV). Cette surestimation est en moyenne d'autant plus forte que la résolution décroît. D'autre part il existe un décalage entre l'échelle de calcul des coefficients en ondelettes $W_T(k,l,a)$ obtenus à partir de l'approximation $T(k,l,\frac{a}{2})$, et l'approximation $T(k,l,a)$. Comme à chaque échelle les approximations $\hat{T}(k,l,\frac{a}{2})$ sont reconstruites à partir d'une combinaison linéaire des coefficients $W_{\hat{T}}(k,l,a)$ et des approximations $\hat{T}(k,l,a)$, elles sont construites à partir d'approximations $\hat{T}(k,l,a)$ dont l'amplitude est surestimée par rapport à l'amplitude des approximations ayant conduit au calcul des coefficients. Ainsi les variations d'émissivité brisent la conservation du flux du paramètre T entre les échelles. Ce défaut est dû à la nature non linéaire de la transformée construite indépendamment d'échelles en échelles. Il est d'autant plus fort que les éléments de résolution incluent des régions de températures distinctes avec un fort gradient d'émissivité.

Simulation numérique à trois dimensions d'un amas en cours de coalescence.

Les figures 9 et 10 de l'article de la section 4.4 illustrent l'observation simulée d'un amas de galaxies, lui-même sélectionné dans une simulation numérique de formation des structures pour être le siège de processus de coalescence violents (Teyssier, 2002, voir aussi section 4.2.2). La

figure 9 montre les distributions d'émissivité et de température projetées du plasma intra-amas d'un amas simulé à trois étapes de son processus de formation. Les âges de l'amas au cours de ces trois étapes correspondent à des redshifts de 0.13, 0.09 et 0.0, respectivement. On observe la fusion de sous-groupes majeurs et d'autres sous structures en un seul amas à $z=0.0$. Les processus de chocs engendrent une élévation de la température globale de l'amas dans sa région centrale par rapport aux températures initiale des sous-groupes. Une zone de compression chaude se développe entre les deux sous-groupes à $z=0.09$, puis des ondes de chocs se dispersent radialement après la fusion à $z=0.0$, conduisant à une structuration en température très perturbée du plasma.

Pour simuler une observation, l'amas a été replacé à une distance correspondant à un redshift de 0.2, de sorte qu'il devient visible dans son ensemble dans le champ de 30 minutes des caméras EPIC. Son émissivité correspond à celle que l'on obtiendrait avec une observation de 60 ks typique d'une demande de temps "longue". Les listes d'événements ont été constituées par J.-L. Sauvageot (CEA), à partir d'un algorithme de type Monte-Carlo qui retrace le parcours de chaque photon depuis une cellule de la simulation numérique jusqu'au plan du détecteur. Ce parcours prend en particulier en compte la déflexion des photons sur les miroirs en fonction de leur énergie, conduisant à une distribution spectrale $S(e)$ qui suit la réponse instrumentale. Une composante de fond particules $P(e)$ uniforme a ensuite été ajoutée au modèle. L'intérêt de cette simulation est de se rapprocher au maximum des conditions d'observations réelles, notamment en modélisant les effets de projection qui conduisent à un spectre d'émission thermique multi-composantes. Si la qualité des images est dégradée par la statistique limitée, on constate qualitativement un bon accord entre les cartes en températures restaurées par l'algorithme de spectro-imagerie et les distributions de température de la simulation projetées sur le plan du ciel. L'algorithme adaptatif permet de bien résoudre les structures proches du centre des sous-groupes où l'émissivité est la plus élevée, mais pas les structures des bords de champ où la statistique est trop faible. Les structures caractéristiques des régions de chocs ainsi que les températures absolues sont bien restituées malgré les effets de projection, ce qui est encourageant quant à la capacité de cet algorithme de mettre en évidence les structurations de température caractéristiques des processus de fusion dans les amas réels.

4.4 Publication à propos de l'algorithme de spectro-imagerie.

Temperature map computation for X-ray clusters of galaxies

H. Bourdin¹, J.-L. Sauvageot², E. Slezak¹, A. Bijaoui¹, and R. Teyssier^{2,3}

¹ Observatoire de la Côte d'Azur, BP 4229, 06304 Nice Cedex 4, France

² CEA/DSM/DAPNIA, Service d'Astrophysique, CEA/Saclay, 91191 Gif-sur-Yvette Cedex, France

³ NIC Consortium (Numerical Investigations in Cosmology, CEA Saclay)

Received 16 April 2003 / Accepted 19 September 2003

Abstract. Recent numerical simulations have shown that the variations of the gas temperature in clusters of galaxies are indicative of the dynamical state of these clusters. Maps of the temperature variation show complex structures with different shapes at different spatial scales, such as hot compression regions, filaments, cooling flows, or large-scale temperature profiles. A new multiscale spectro-imagery algorithm for restoring the spatial temperature variations within clusters of galaxies is presented here. It has been especially developed to work with the EPIC MOS1, MOS2 and PN spectro-imagers on board the XMM-Newton satellite. The temperature values are fitted to an emission model that includes the source, the cosmic X-ray background and cosmic-ray induced particle background. The spatial temperature variations are coded at different scales in the wavelet space using the Haar wavelet and denoised by thresholding the wavelet coefficients. Our local temperature estimator behaves asymptotically like an optimal minimum variance bound estimator. But it is highly sensitive to the instrumental and astrophysical backgrounds, so that a good knowledge of each component of the emission model is required. Our algorithm has been applied to a simulated 60 ks observation of a merging cluster at $z = 0.1$. The cluster at different stages of merging has been provided by 3-D hydrodynamical simulations of structure formation (AMR). The multiscale approach has enabled us to restore the faint structures within the core of the merging subgroups where the gas emissivity is high, but also the temperature decrease at large scale in their external regions.

Key words. galaxies: clusters: general – galaxies: interactions

1. Introduction

Within the context of the so-called Λ -Cold Dark Matter scenario for hierarchical formation of structures, clusters of galaxies are the largest structures and consequently the last ones to have been formed. They are supposed to have undergone recently the transition between linear and non-linear regimes. Actually, clusters of galaxies are formed by aggregation of groups and subclusters of different masses along filamentary over-densities, as shown by numerical simulations (Katz & White 1993; Colberg et al. 2000; Evrard 1990; Evrard et al. 2002).

Most of the mass of clusters of galaxies is split into a dark matter component – around 70% of the total mass – and a gas halo component. The gas component is an optically thin plasma heated up to 10^8 K and confined within the gravitational potential at thermal equilibrium, which entails a thermal bremsstrahlung emission detectable in X-rays. Since the cluster accretion process is mainly driven by gravitation, observing the X-ray emission of this gas provides a significant piece of information about the dynamical state of the cluster as a whole. Besides, such studies are only limited by photon statistics since

the intra-cluster gas is an extended source, whereas optical studies using galaxy counts are limited by the galaxy distribution. Recent hydrodynamic simulations (Ritchie & Thomas 2001; Ricker & Sarazin 2001; Roettiger et al. 1996; Teyssier 2002) have shown the sensibility of the gas surface brightness and temperature variations to events related to post and on-going mergers such as gas compressions, cooling flows and cold fronts. But our understanding of the details of the merger scenario, especially the physical processes occurring in the dissipative phase, must strongly be improved by benefiting from new observational data on nearby clusters. The new generation of X-ray telescopes (Chandra; XMM-Newton) now provides data with enough spatial and spectral resolution to try to detect such variations. However, the X-ray photon statistics with its typical emission of a few photons per second in a field of view (FOV) of 30 arcmin in diameter leads to a much lower signal-to-noise ratio than obtained with usual optical observations. Consequently any analysis of X-ray spectroscopic data must involve a spatial binning to increase the photon statistics. To do so, while preserving the best spatial resolution, multiscale algorithms can be applied. We introduce here such a multiscale spectro-imagery algorithm dedicated to the optimal mapping of temperature variations in the observed intra-cluster plasma.

Send offprint requests to: H. Bourdin,
e-mail: bourdin@obs-nice.fr

This algorithm has been especially designed to process data from the XMM-EPIC spectro-imagers, MOS1, MOS2 and PN cameras. Thanks to a maximum likelihood algorithm the temperature of the gas inside each element of a grid covering the field of view is computed with its signal-to-noise ratio. Several computations are done for grids with different spatial resolution following a dyadic scheme, in order to obtain the final temperature map from significant temperature variations at different scales. To do so a wavelet formalism is applied: a selection of significant wavelet coefficients at different scales enables us to reconstruct the optimal temperature map. The spectral analysis is introduced in the first section while the multiscale algorithm is described in the second one. Results from simulations of EPIC-XMM data on merging clusters are discussed in the third section. The last section is devoted to our conclusions.

2. Local temperature computation: Spectral analysis

The EPIC spectro-imagers work in photon count mode. So, the data to be processed are a set of photon events, composed of the photon energy and its location on the detector surface. Given the events within an area of the detector we get the observed emission spectrum of the related region of the sky. By modeling this emission into its different astrophysical and instrumental components, the proper cluster emission can be inferred from this spectrum and the temperature and emissivity of the gas can be derived using an optically thin plasma emissivity model.

2.1. The emission model

The observed emission spectrum has been modelled as a combination of three different emission components. The first of these components is the proper cluster emission, the second one is the cosmic X-ray background (CXB) and the last one is a “particle background” which has been introduced to model the EPIC detector response to induced particles from interactions of cosmic particles with the satellite platform.

The first two components correspond to astrophysical sources. In that case the photons have to follow the telescope optical path before reaching the detector and the emission spectra of these sources must be convolved with the spectral instrumental response. The EPIC XMM-Newton photon counts suffer from a vignetting effect which depends both on coordinates on the detector and on the photon energy. Considering this energy dependence, the vignetting function of each EPIC camera is more or less flat at low energy (0.5–5 keV), but it decreases at high energy (>5 keV), so that the net statistics are poorer at high energy. The spatial vignetting function decreases the photons counts in the focal plane as a function of the distance to the detector optical axis. This vignetting effect has to be taken into account when considering the spectral instrumental response. The convolution of the CXB and cluster emissivity component can be done for each camera using the associated vignetting function provided that the mean spatial location of the analyzed events on the detector surface is known.

The cluster X-ray emission model is modelled as an optically thin plasma emission (Mewe 1985, 1986). The plasma

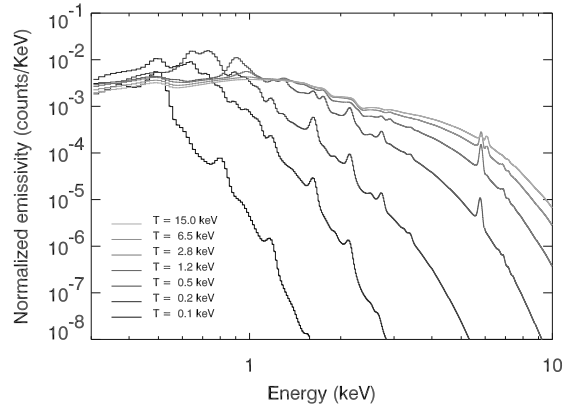


Fig. 1. Normalized emissivity $S(T, e)$ of the optically thin plasma as a function of temperature, ($z = 0.1$, $N_H = 6 \times 10^{20} \text{ cm}^{-2}$). The spectra have been convolved by the EPIC-XMM MOS1 focus instrumental response.

is considered to be a hot ($T > 10^6 \text{ K}$) and low-density one ($n \approx 10^{-3} \text{ cm}^{-3}$) at the coronal limit (Sarazin 1986). Assuming the plasma to be at the ionization equilibrium this emission model is composed of two components. The first component is a continuum of free-free and free-bound radiation emitted during the Coulomb scattering of the electrons by the plasma ions. The second one is a line radiation due to recombination of electrons after the collision process of the different elements of the plasma. Under these assumptions the shape of the emission spectra for the galaxy cluster only depends on the plasma equilibrium temperature T and elements abundances. But the optical path from the source to the detector entails three distortions of this spectral shape when observed: the redshift, the X-ray absorption due to ionization of galactic hydrogen and the EPIC-XMM instrumental response. So, given the element abundances of the plasma, the redshift of the source, the hydrogen column density between the observer and the source and the EPIC-XMM instrumental response, a table of normalized emission spectra $S(T, e)$ can be computed with temperature T and energy e . Such a table is plotted in Fig. 1, where the emission spectra have been convolved with the EPIC-XMM MOS1 focus instrumental response, where no vignetting occurs.

The cosmic X-ray background (CXB) is composed of three components. The first two components correspond to the “soft X-ray background” due to the hot local bubble which can be modelled as an optically thin plasma emission at 0.204 and 0.074 keV respectively. The third component is modelled using a power-law emission with index $\gamma = 1.42$ and corresponds to the hard X-ray background (Lumb et al. 2002). As for the cluster component, the CXB component have to be convolved with the XMM-EPIC instrumental response. Figure 2 shows the three X-ray background spectra convolved with the EPIC-XMM focus instrumental response.

For each camera, the “particle background” can be modelled as a flat energy response plus some fluorescence lines of the detector in the 1.4–2 keV band (Al, Si). Contrary to the CXB the particles background does not correspond to photons

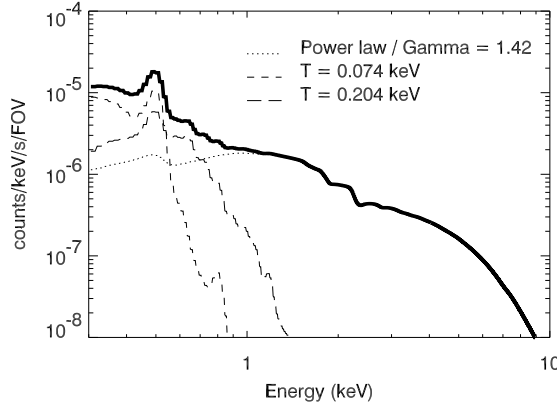


Fig. 2. Three component X-ray background $N_B B(E)$. The three components have been convolved with the EPIC-XMM MOS1 focus instrumental response.

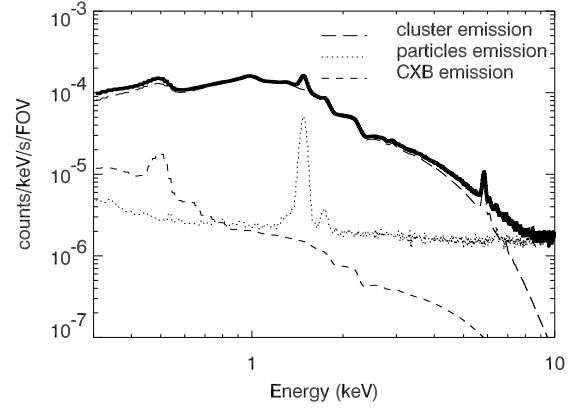


Fig. 3. EPIC-XMM MOS1 total focus emissivity in the line of sight of an isothermal cluster at 5 keV, taking into account the cluster component, but also CXB and particle background.

following the telescope optical path but to particles impacting the CCD without interacting with the telescope mirrors. Consequently the particle background is supposed to be uniform on the detector surface and needs not be convolved with the instrumental response.

Let us call $P(e)$ and $B(e)$ the normalized particles and astrophysical background components, respectively. We get an emission spectrum with three components, where the relative contribution of each component, i.e. $S(T, e)$ for the cluster, $B(e)$ for the CXB and $P(e)$ for the particles, can be taken into account by means of three normalization terms, N_S , N_B and N_P , respectively. Let us call $F(T, e)$ the normalized emission model and N_F its normalization term, we get:

$$N_F F(T, e) = N_S S(T, e) + N_B B(e) + N_P P(e), \quad (1)$$

where the normalization terms are linked to each other:

$$N_F = N_S + N_B + N_P. \quad (2)$$

Figure 3 shows the on focus EPIC-XMM MOS1 total emissivity $N_F F(T, e)$ of a simulated isothermal cluster at 5 keV at redshift $z = 0.1$ and for $N_H = 6 \times 10^{20} \text{ cm}^{-2}$, and the related three components, $N_S S(T, e)$, $N_B B(e)$ and $N_P P(e)$.

2.2. The maximum likelihood function

Let us now consider a set of N events of known location on the detector surface corresponding to a simulated or real galaxy cluster observation. Providing that the mean background emissivities N_P and N_B are known, the temperature T and emissivity N_S of the cluster are the unknown parameters to be fitted to the global emission model $F(T, e)$.

The mean background emissivities N_P and N_B are supposed to be known and uniform over the detector surface. Indeed, considering that events in the 10–12 keV energy band are particles events exclusively, the value of N_P can be computed by adding all the events within the 10–12 keV range and extrapolating the bolometric particle emissivity from the particle emission model. The value of N_B can be estimated empirically from

the emissivity of external regions of the FOV where the CXB emissivity overcomes the cluster emissivity. Introducing the emissivity parameter $r = \frac{N_S}{N_F}$ to get a bounded parameter space, Eqs. (1) and (2) can be rewritten as:

$$\begin{aligned} F(r, T, e) &= \left[1 - \frac{r}{N} (N_B + N_P) \right] S(T, e) \\ &\quad + \frac{r}{N} [N_B B(e) + N_P P(e)], \\ &= S(T, e) + \frac{r}{N} [N_B (B(e) - S(T, e)) \\ &\quad + N_P (P(e) - S(T, e))]. \end{aligned} \quad (3)$$

$$N_S = N_F - N_B - N_P,$$

$$= N \left[\frac{1}{r} - \frac{1}{N} (N_B + N_P) \right]. \quad (4)$$

So a two parameter model fitting algorithm has to be used to determine simultaneously r and T . A pursuit of the maximum of the likelihood function has been chosen because, contrary to the mean squares estimator, this estimator gives significant results, even in cases of bad statistics. To do so, the emission model is sampled in energy. Each event i corresponds to a photon detection in the energy bin e_i , so that the logarithm of the likelihood function (Kendall & Stuart 1973) can be written as:

$$\ln [L(r, T)] = \sum_i \ln [F(r, T, e_i)]. \quad (5)$$

The maximum of $\ln [L(r, T)]$ gives the values of both parameters T and r . This maximum has to be searched for within a bounded parameter space. Indeed, the emission model $F(r, T)$ only makes sense when both parameters r and T are positive. Furthermore the emission spectra are modelled in a temperature interval corresponding to the intra-cluster plasma usual temperatures, while the emissivity parameter r must be close to 1. Consequently, the parameter space is usually bounded in temperature, from $T_{\min} = 0.1 \text{ keV}$ to $T_{\max} = 30 \text{ keV}$, but also in the emissivity parameter, whose bounds have been set to

$r_{\min} = 0$ and $r_{\max} = 2^1$. In order to pursue the maximum of $\ln[L(r, T)]$ within such a bounded parameter space a modified conjugate gradient algorithm has been chosen for its fast convergence (cf. Appendix A). Eventually, given both values of T and r , the error-bars σ_T and σ_r are obtained from the diagonal elements of the matrix V , deduced from the Fisher matrix F_M :

$$V = -[F_M]^{-1}, \quad \text{with} \quad F_M = \begin{bmatrix} F_1 & \rho \\ \rho & F_2 \end{bmatrix},$$

where:

$$F_1 = \frac{\partial^2 \ln(L)}{\partial r^2}, \quad F_2 = \frac{\partial^2 \ln(L)}{\partial T^2}, \quad \rho = \frac{\partial^2 \ln(L)}{\partial T \partial r}.$$

3. Global detection of temperature variations: A multiscale algorithm

3.1. Introduction

Let us spatially sample the EPIC-XMM FOV along both axes k and l . Given an event-list associated with the pixel k, l , but also an emission model $N_F F(T, e, k, l)$ taking into account the local instrumental response, the maximum likelihood method described in the previous section enables us to define a local estimator $\widehat{T}^{k,l}$ of the parameter $T(k, l)$.

To compute the $\widehat{T}^{k,l}$ value with a high enough signal-to-noise ratio the statistics may have to be enhanced by means of a spatial binning, but this decreases the spatial resolution. Consequently, given the requested signal-to-noise value, an optimal local bin size has to be computed to define $\widehat{T}^{k,l}$, which entails a multiscale exploration of the $\widehat{T}^{k,l}$ signal. A natural multiscale exploration of a given signal $I(k, l)$ is provided by the wavelet transform $W(a, k, l)$ (cf. Appendix B), where the signal $I(k, l)$ is unfolded in the space-scale domain, with a as the scale parameter. Let us define $W_{\widehat{T}^{k,l}}(a)$ the wavelet transform of $\widehat{T}^{k,l}$. The aim of our algorithm is to show the significant $\widehat{T}^{k,l}$ variations at different scales thanks to a thresholding of the wavelet space.

The chosen wavelet transform is based on the Haar wavelet as will be discussed in Sect. 3.2. As a first step, successive binnings of the EPIC-XMM data set at each scale and fitting of the emission model $N_F F(T, e, k, l)$ of Eq. (2) with the maximum likelihood algorithm enable to define successive approximations, $A_{\widehat{T}^{k,l}}(a, k, l)$, of the $\widehat{T}^{k,l}$ distribution. Then the “first guess” set of wavelet coefficients $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$ is computed by applying the high-pass analysis filters corresponding to the Haar wavelet to $A_{\widehat{T}^{k,l}}(a, k, l)$. Notice that this set of coefficients does not constitute a real wavelet transform since the coefficients are computed by means of different emission models from scale to

scale. Nevertheless, it may be considered as an approximation of the real wavelet transform $W_{\widehat{T}^{k,l}}(a, k, l)$ that we are looking for. Within the wavelet space the respective noise and signal contribution to the signal can be easily separated at each scale. This will be described for the chosen transform in Sect. 3.3. The last step, developed in Sect. 3.4, consists of removing iteratively the noise contribution to $W_{\widehat{T}^{k,l}}^{(0)}$, which leads to a new set of denoised wavelet coefficients $W_{\widehat{T}^{k,l}}^{(n)}(a)$. The final set of coefficients will consist of the expected $W_{\widehat{T}^{k,l}}(a)$ wavelet transform. This process is performed from the lowest to the highest resolution allowing us to construct in the direct space better and better approximations of $\widehat{T}^{k,l}$ and to obtain the optimal temperature map $\widehat{T}(k, l)$.

3.2. Computing the $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$ set of coefficients

So as to compute the set of coefficients $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$ a wavelet function has to be chosen. The 2-D Haar wavelet (Haar 1910) which matches the square spatial sampling is the most suitable since a square binning has been chosen. It is a compact support square kernel with coefficients $\frac{1}{2}$ and $-\frac{1}{2}$, following the amplitude normalization condition of the wavelet coefficients. The Haar wavelet is anisotropic, so that the wavelet transform results at each scale $a = 2^j$ with $j = 0, \dots, N$ in a decomposition onto three detail spaces with wavelet coefficients $W_p(a, k, l)$ where $p = h, v, d$ indicates the horizontal, vertical and diagonal details, respectively. Their value can be computed from the convolution of the approximation at scale $\frac{a}{2}$, $A(\frac{a}{2}, k, l)$, using the high-pass analysis filter $g_p(\frac{a}{2})$ (cf. Appendix B).

Let us associate with each pixel k, l of the FOV the set of events found within a square bin of size $\frac{a}{2}$. The fitting of this set of events to the model $N_F F(T, e, \frac{a}{2}, k, l)^2$ will enable us to compute the set of coefficients $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)^3$. To do so, the approximations $A_{\widehat{T}^{k,l}}^{(0)}(\frac{a}{2}, k, l)$ with their rms values $\sigma_{A_{\widehat{T}^{k,l}}^{(0)}}(\frac{a}{2}, k, l)$ are first computed thanks to the maximum likelihood fitting of the model $N_F F(T, e, \frac{a}{2}, k, l)$ to the data corresponding to pixel k, l with scale $\frac{a}{2}$. Then the wavelet coefficients $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$ are computed by applying the high-pass analysis filters $g_p(\frac{a}{2})$ (cf.

Fig. B.2). The variances $[\sigma_{W_{\widehat{T}^{k,l}}^{(0)}}(a, k, l)]^2$ of the wavelet coefficients $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$ are computed by applying the low-pass analysis filter $h(\frac{a}{2})$ to the variances of the smoothed plane $[\sigma_{A_{\widehat{T}^{k,l}}^{(0)}}(\frac{a}{2}, k, l)]^2$.

The binning of the data has been chosen such that the smoothed planes $A_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$ have the same number of pixels whatever the scale is, yielding a redundant wavelet transform. As it will be explained in the following section the redundancy

¹ Considering an emission model with a given emissivity N_F , the number of observed counts N is expected to fluctuate like the standard deviation of the Poisson parameter N_F : $\sigma_{N_F} = \sqrt{N_F}$, so that the $r = \frac{N}{N_F}$ values are expected to fluctuate like $\sigma_r = \frac{\sqrt{N_F}}{N_F}$. Consequently the expected r values can be bounded within the $[-\Delta_r, +\Delta_r]$ interval, provided that $\Delta_r > \sigma_r$. The Δ_r value of 1, corresponding to σ_r for one photon per pixel was chosen. Indeed, σ_r decreases with increasing statistics, – for instance $\sigma_r = 0.32$ for $N_F = 10$ –, so that we always get $\sigma_r < \Delta_r$.

² The scale parameter a has now been added because $N_F F(T, e, \frac{a}{2}, k, l)$ is in fact the weighted mean of emission models corresponding to each event found within the square bin of edge a .

³ The p index is hereafter omitted because the following processes do not depend on the orientation of the coefficients.

has been chosen in order to balance the usual poor statistics of the data.

3.3. The noise contribution: The wavelet coefficients thresholding

We get a set of coefficients $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$ with zero mean value and variances $\sigma_{W_{\widehat{T}^{k,l}}^{(0)}}(a, k, l)$. Given the signal-to-noise ratio R , the coefficients $W_{\widehat{T}^{k,l}}^{(0),s}(a, k, l)$ with value overcoming R times the local variance will be considered as significant whereas the other ones $W_{\widehat{T}^{k,l}}^{(0),ns}(a, k, l)$ will be considered as non-significant:

$$\begin{aligned} |W_{\widehat{T}^{k,l}}^{(0),s}(a, k, l)| &> R \sigma_{W_{\widehat{T}^{k,l}}^{(0)}}(a, k, l), \\ |W_{\widehat{T}^{k,l}}^{(0),ns}(a, k, l)| &< R \sigma_{W_{\widehat{T}^{k,l}}^{(0)}}(a, k, l). \end{aligned} \quad (6)$$

Let us attribute the Boolean true value to regions of the wavelet space where the $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$ coefficients have been considered as statistically meaningful, and conversely the Boolean false value to regions where they have been considered as non-meaningful. Since the set of coefficients $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$ is an approximation of the expected wavelet transform $W_{\widehat{T}^{k,l}}(a, k, l)$, the projection of this Boolean space onto the successive wavelet planes at different scales defines an approximation of the so-called significant masks $M(a, k, l)$ of the genuine wavelet transform $W_{\widehat{T}^{k,l}}(a, k, l)$.

Notice that due to the redundancy of the algorithm, the wavelet coefficients $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$ are computed from dependent data. A crude estimation of dependency of the coefficients of a wavelet plane is that they are linked to each other, providing that they are included within a region of area a^2 . This information has been used to remove some false detections of significant regions. The connected significant regions of $M(a, k, l)$ are kept providing that their area overcomes a^2 , but are considered non-significant otherwise.

3.4. Constructing the $W_{\widehat{T}^{k,l}}(a, k, l)$ wavelet transform and temperature map $\widehat{T}(k, l)$

This splitting of the wavelet space enables us to construct a true wavelet transform $W_{\widehat{T}^{k,l}}(a, k, l)$ from $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$. Since all the available information about the variations of parameter $\widehat{T}^{k,l}$ is considered as being contained within the significant domain of the wavelet space, the wavelet transform $W_{\widehat{T}^{k,l}}(a, k, l)$ can be constructed by keeping the coefficient values of the significant domain of $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$. Conversely the coefficient values of the non-significant domain, where information about the variations of $A_{\widehat{T}^{k,l}}^{(0)}(\frac{a}{2}, k, l)$ is considered as lost, will be iteratively modified. Since no variation of $\widehat{T}^{k,l}(k, l)$ has been detected within this domain, a priori information will be added in place of the lack of information: it will be supposed that the $\widehat{T}^{k,l}$ variations are minimal. This condition leads to the Tikhonov smoothness constraint or regularization, where the gradient of the smoothed

planes defined as $\|\frac{\partial A}{\partial x}\|^2 + \|\frac{\partial A}{\partial y}\|^2$ have to be minimized, which implies the following condition (cf. Bobichon & Bijaoui 1997):

$$L_2 A_{\widehat{T}^{k,l}}^{(n)}\left(\frac{a}{2}, k, l\right) = 0, \quad (7)$$

where L_2 is the 2-D Laplacian operator.

Let us call $w_{A_{\widehat{T}^{k,l}}^{(n+1)}}(a, k, l)$ and $w_{L_2 A_{\widehat{T}^{k,l}}^{(n)}}(a, k, l)$ the wavelet coefficients computed by applying the high pass analysis filters $g_p(\frac{a}{2})$ to the respective approximations, $A_{\widehat{T}^{k,l}}^{(n+1)}(\frac{a}{2}, k, l)$ and $A_{\widehat{T}^{k,l}}^{(n)}(\frac{a}{2}, k, l)$, Eq. (7) is solved thanks to an iterative Van-Cittert algorithm on the wavelet coefficients constructed by successive filtering and reconstruction of the smoothed planes $A_{\widehat{T}^{k,l}}^{(n)}(\frac{a}{2}, k, l)$ at scale $\frac{a}{2}$:

$$w_{A_{\widehat{T}^{k,l}}^{(n+1)}}(a, k, l) = w_{A_{\widehat{T}^{k,l}}^{(n)}}(a, k, l) - \beta w_{L_2 A_{\widehat{T}^{k,l}}^{(n)}}(a, k, l). \quad (8)$$

Since $w_{A_{\widehat{T}^{k,l}}^{(n+1)}}(a, k, l)$ and $w_{L_2 A_{\widehat{T}^{k,l}}^{(n)}}(a, k, l)$ correspond to the wavelet coefficients of the temperature map $\widehat{T}(k, l)$ at scale a , Eq. (8) can be rewritten as:

$$W_{\widehat{T}^{k,l}}^{(n+1)}(a, k, l) = W_{\widehat{T}^{k,l}}^{(n)}(a, k, l) - \beta W_{L_2 A_{\widehat{T}^{k,l}}^{(n)}}(a, k, l), \quad (9)$$

where β is a fixed convergence parameter. An additional condition prevents the non-significant coefficient values from overcoming the significance threshold. The convergence criterion $\|W^{(n)} - W^{(n+1)}\| < \varepsilon$, is reached in a few iterations with $\varepsilon = 0.01$. This construction of the wavelet transform $W_{\widehat{T}^{k,l}}(a, k, l)$ is done from the coarsest resolution to the finest one while constructing successive smoothed planes $A_{\widehat{T}^{k,l}}^{(n)}(\frac{a}{2}, k, l)$. The final smoothed plane at the finest resolution defines the final temperature map $\widehat{T}^{k,l}(k, l)$ we are looking for. Notice that the final wavelet transform $W_{\widehat{T}^{k,l}}(a, k, l)$ is the shift invariant wavelet transform introduced by Coifman & Donoho (1995).

3.5. Error map computation

We describe the coarse method we adopted to estimate the final error on the temperature map $\widehat{T}^{k,l}(k, l)$. The temperature map $\widehat{T}^{k,l}(k, l)$ can be considered in the direct space as a linear combination of the different wavelet coefficients $W_{\widehat{T}^{k,l}}^{(n)}(a, k, l)$. For each pixel k, l , the $\widehat{T}^{k,l}(k, l)$ standard deviation is mainly due to the wavelet plane $W_{\widehat{T}^{k,l}}^{(n)}(a, k, l)$ standard deviation, where the scale a corresponds to the highest resolution where $W_{\widehat{T}^{k,l}}^{(n)}(a, k, l)$ is found significant. Consequently, we decided to estimate the error-bar $\sigma_{\widehat{T}^{k,l}}(k, l)$ from the $W_{\widehat{T}^{k,l}}^{(n)}(a, k, l)$ standard deviation. Since $W_{\widehat{T}^{k,l}}^{(n)}(a, k, l)$ is included within the significant domain, the $W_{\widehat{T}^{k,l}}^{(n)}(a, k, l)$ standard deviation is close to the $W_{\widehat{T}^{k,l}}^{(0)}(a, k, l)$ standard deviation value, $\sigma_{W_{\widehat{T}^{k,l}}^{(0)}}(a, k, l)$. We get at pixel k, l ,

$$\sigma_{\widehat{T}^{k,l}}(k, l) = \sqrt{\sum_p \left[\sigma_{W_{\widehat{T}^{k,l}}^{(0)}}(a, k, l) \right]^2}, \text{ where } a \text{ is the highest resolution of structure detection.}$$

4. Algorithm performances against simulations

The performance and limits of the algorithm have been tested by means of a set of simulated observations with different

values of source temperatures and emissivities, as well as background emissivities. The spectral fitting performance will be introduced in a first section with uniform source temperatures and emissivities. Then the multiscale algorithm will be tested in a second section on model EPIC-XMM observations with non-uniform temperatures and emissivities but with only one gas emission component $F(k, l, e)$.

4.1. Spectral fitting

4.1.1. Introduction

The spectral fitting performance of our algorithm has been tested against different values of the parameters of the emission model – cf. Eq. (3) –, taking or not taking into account instrumental components. To do so, a cube $F(k, l, e)$ is constructed for different values of $T(k, l)$ and $N_S(k, l)$, the temperature and emissivity of the source, respectively, as well as for N_B and N_P , the background emissivities. The chosen simulated camera is the EPIC-XMM MOS1. An observation can be simulated by drawing randomly within the Poisson distribution of parameter $F(k, l, e)$ a value for the number of photons at position k, l, e . The models used hereafter are 128×128 square images where the source temperature and emissivity are considered uniform, and fixed at 3 keV and 100 photons per pixel, respectively. The fitting is first tested without any background emissivity and then with different particles and CXB background emissivities values. In each case the vignetting influence was also tested by constructing emission models with or without vignetting. Just like for the first step of our multiscale algorithm, the fitting of $F(k, l, a)$ is performed at different resolutions a , leading to different maps $T(k, l, a) \equiv A_{\text{PQ}}^{(0)}(a, k, l)$ with $a = 2^j$, $j = 0 \dots N$. Actually the fitting has been performed on four successive resolutions with two emission models at 100 and 200 photons per pixel respectively, leading to eight maps with statistics increasing by a factor of two, from 200 to 25 600 cluster photons per pixel. The first spectral fitting test is performed without any background. The mean, standard deviations and a selection of six histograms of the different maps $T(k, l, a)$ corresponding to this spectral fitting are plotted in Fig. 4. An instrumental response with vignetting factor was introduced for the first histogram, but without vignetting for the second one. The histograms of Figs. 5 and 6 correspond to emission models with particles and CXB background, respectively, where each background emissivity values have been set to 0.5 and 2 times the value of the cluster emissivity.

4.1.2. No background vignetted emission model

Let us first consider the fitting of the isothermal emission model $F(k, l, e)$ at 3 keV with a null background emissivity (cf. Fig. 4). The temperature distributions obtained from fittings with more than 1600 cluster photons per pixel are symmetric around a peak value at 3 keV.

The maximum likelihood method used for the model fitting should provide a consistent estimator whose variance is close to the Rao-Cramer limit (Kendall & Stuart 1973) of the Minimum Variance Bound (MVB) estimators. We have checked that the

variances of the temperature histograms are close to the mean of the MVB variances values computed from the Fisher matrix after each spectral fitting. They follow the statistics with N counts, so that $\sigma \propto \frac{1}{\sqrt{N}}$. The variances of distributions obtained from fittings with more than 400 cluster photons per pixel are in agreement with the MVB variance, but the variance of the distribution obtained with 200 cluster photons per pixel is not.

The dispersion obtained with 200 cluster photons is asymmetric and does not peak at the 3 keV value, so that a bias is evidenced. This “low temperature bias” is due to the emissivity of the high-energy part of the source emission spectra, with a variation of several orders of magnitude between the high and low temperature emission models (cf. Fig. 1), and a high relative chance fluctuation of the high-energy photons number. It has been reduced by an histogram equalization of the energy distribution of the photons. It is not worth fully removing this bias since it is only important in the case of very poor statistics when the temperature dispersion is high anyway. Instead of removing it, two conditions have been added to the spectral fitting. The result of the temperature fit of the different resolution elements is only taken into account providing that on the one hand, the 1σ error bars values are lower than 1 keV and on the other hand, more than about 500 photons are available so that the statistics is considered high enough.

4.1.3. No background non-vignetted emission model

Now let us recall that due to the spatial variations of the detector response the different temperatures values are not computed following the same process which is questionable. To get rid of this drawback, but also to test the vignetting influence on the spectral fitting, a model with the same source component as the previous one but without vignetting has been built (cf. Fig. 4). The detector response is now uniform and similar to the focus response all over the detector surface. We get a close to MVB estimator at each resolution with lower temperature dispersions than the dispersions obtained when the vignetting effect occurs. Indeed, the removal of the vignetting effect results in providing much higher energy photons, which enhances the statistics and reduces the “low temperature bias”.

4.1.4. Vignetted emission model with particles background

The particles background influence on the spectral fitting is tested thanks to a model series with vignetting and the same source temperature and emissivities as in Sect. 4.1.1 but with particles background emissivities of 0.5 and 2 times the cluster emissivity (cf. Fig. 5). The temperature variance is larger than in the case of null background emissivity and all the larger since the background emissivity increases. The fitted source temperatures are found to be still distributed around the $T = 3$ keV value in the case of a cluster emissivity higher than 3200 photons per pixels, but biased in the case of lower statistics. The “low-temperature bias” is enhanced here since the source high-energy photons are much less numerous than

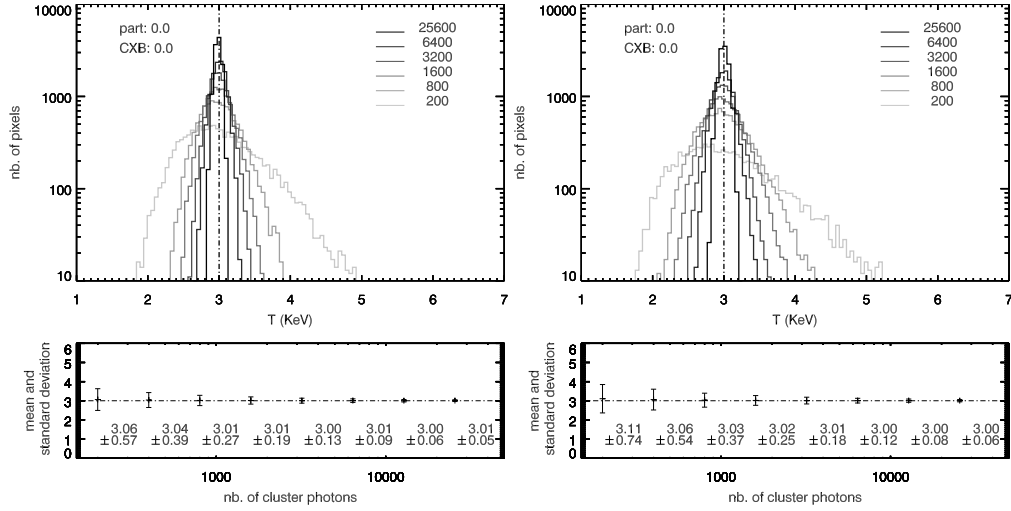


Fig. 4. Distribution of the results of the temperature fits for a simulated EPIC-XMM MOS1 observation of an isothermal extended source. The source emissivity is uniform. No background emission is included. The histograms with increasing darkness correspond to three fixed source emissivity N_S , from $N_S = 200$ to 25 600 photons per pixel. *Left*: fittings performed with uniform instrumental response simulated. *Right*: fittings performed with real instrumental response simulated.

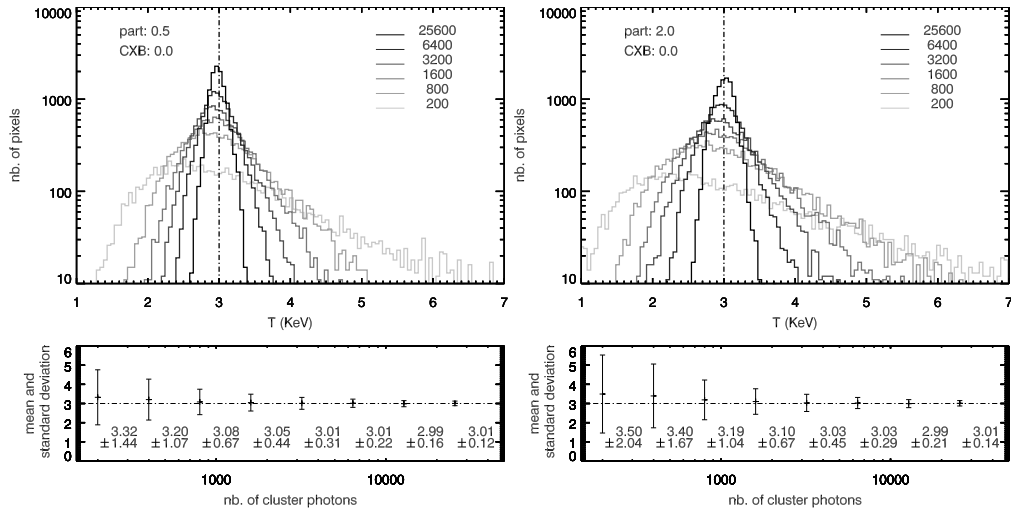


Fig. 5. Same as Fig. 4 with real instrumental response and particles background emission simulated. *Left*: background emissivity value of 0.5 times the cluster emissivity value. *Right*: Background emissivity value of two times the cluster emissivity value.

the high-energy particles. Yet despite the bias, the temperature estimator follows the MVB variance in the case of a cluster emissivity value higher than 400 photons per pixel.

4.1.5. Vignetted emission model with CXB

Let us now discuss the CXB influence on the spectral fitting thanks to a model series with vignetting and CXB emissivities of 0.5 and 2 times the cluster emissivity (cf. Fig. 6). Whatever the CXB emissivity is, the fitted temperatures peak at 3 keV

in the case of a cluster emissivity value higher than 1600 photons per pixel, with lower dispersion and temperature bias than for the toy-model series with particles background. Also, the temperature estimator behaves like an MVB estimator. The vignetted CXB provides many fewer high-energy photons than the particles background. Consequently, on the one hand the global background emission is lowered which reduces the temperature dispersion and on the other hand the lack of background low-energy photons reduces the low-temperature bias.

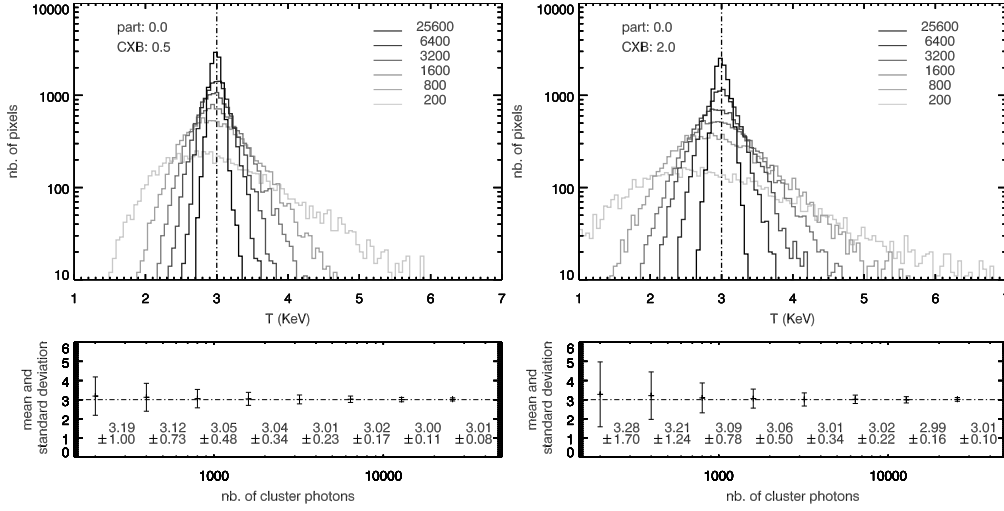


Fig. 6. Same as Fig. 4 with real instrumental response and cosmic X-ray background emission simulated. *Left:* background emissivity of 0.5 times the cluster emissivity value. *Right:* Background emissivity value of two times the cluster emissivity value.

4.1.6. Conclusion

Our temperature estimator is based on the maximum likelihood function computed from the emission model $F(T, e, k, l)$, which provides a close to MVB estimator. Actually the Rao-Cramer limit of MVB estimators cannot be reached, due to the dependency between both parameters r and T in the case of an emission model with background. The variance and bias of the estimator cannot be estimated analytically since the emission model is a numerical model, but tests like those performed enable us to evaluate them for different emission models. The temperature estimator behaves like an MVB estimator in the case of good statistics – more than about 500 photons per pixel – and low background – less than half of the cluster emissivity in our case. Its use in cases of low statistics or high background level is more critical and results must be analyzed carefully.

4.2. Multiscale temperature maps of simulated EPIC-XMM MOS1 observations

Let us now test the second step of our algorithm, i.e. the multiscale analysis, on non-isothermal temperature maps built from a model. The simulated source is an EPIC-XMM MOS1 2D model observation with emission model $F(T, e, k, l)$. Both particles and CXB background emissivities, N_P and N_B , have been set to five photons per pixel. The source is now composed of four isothermal emitting regions: an external ring at 3 keV, two elliptical regions at 4 and 6 keV respectively and the remaining central region at 5 keV. Three emissivity models following a 2-D Gaussian distribution with total flux of 10^5 , 10^6 and 10^7 photons are tested. The corresponding emissivity per pixel range from 10 photons at the edge of the images to 100, 1000 and 10000 photons, respectively, at the centre of the image.

The images size is 128×128 and the wavelet analysis was performed on six scales in order to prevent large scale edge effects, so that details with typical width from 1 to 32 pixels are analysed. The source temperature distribution and reconstructed temperature maps are shown in Fig. 7. The temperature maps have been cut in the middle of each picture so as to show the temperature variations with error-bars in Fig. 8. As expected, the better the statistics is the better the model is reproduced. Due to the shape of the emissivity map of the source (cf. top-left plot in Fig. 8), the temperature variations can only be seen in the central regions with high statistics whereas the external regions with low statistics have been strongly smoothed. The temperature variations of the central regions have been shown at the first scale for the model with 10^7 photons, but have been smoothed due to the lower statistics for both other models. In the external regions, a large-scale mean temperature is computed whatever the global emissivity is due to the low statistics as well as the isothermality of the model.

The quality of the result can be checked by comparing the temperature profiles along the dashed lines of Fig. 7 with the model temperature values (cf. Fig. 8). Notice that such a comparison is not straightforward due to the spatial rebinning of the restored temperature maps. That is why the temperature profile of each restored map needs to be compared with the profile of a “rebinned” model map. The rebinned model maps are computed by applying the wavelet algorithm to the model temperature map with same thresholding of the coefficients as applied for each restored map. To do so, the temperature values are averaged within grids at different scales instead of fitting the data. Two rebinned model map are compared with each restored map. The first one is obtained by computing the arithmetical mean of the temperatures of the model map within the different bins. The second one is obtained by computing a

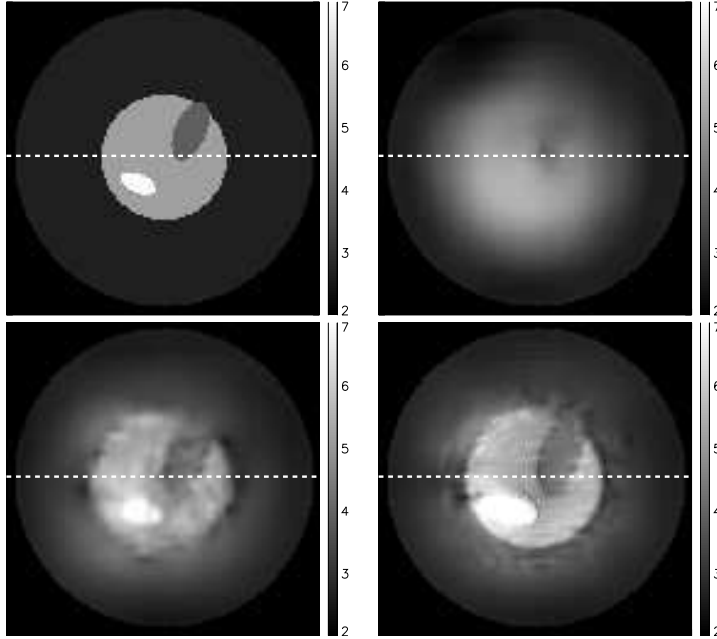


Fig. 7. Temperature maps of simulated EPIC-XMM MOS1 observations, – temperatures coded in keV –, temperature model. *Top-right*: reconstructed temperature map for a global emissivity of 10^5 photons. *Bottom left and right*: reconstructed temperature map for a global emissivity of 10^6 and 10^7 photons, respectively. The dashed line on the reconstructed maps corresponds to profiles displayed in Fig. 8.

temperature mean which has been weighted by the cluster emissivity. The temperature profile of both rebinned model maps are superimposed to each restored map by means of dotted and dashed lines, respectively. Both cuts are closer to the cuts of the restored temperature maps than the cut of the non-rebinned temperature model. The cut computed from weighted means of the temperature model is in agreement with the restored temperature maps, since it is included within the corresponding error-bars.

To conclude about these simulated observations, the multi-scale algorithm has enabled us to adapt the spatial resolution to the local available statistics. The significance of the result has been checked thanks to a multiscale unfolding of the model temperature map. The agreement between the rebinned model map and restored map shows that the spectral fitting is satisfactory whatever the scale or position where it is performed, despite the spatial variation of the likelihood function.

5. Temperature maps of numerically simulated merging clusters

5.1. Hydrodynamic simulation of a merging cluster

In order to test the ability of our algorithm to detect the temperature variations expected within merging clusters, the observation of a cluster at different stages of merging have been simulated. This cluster has been obtained from adaptive mesh refinement 3-D hydrodynamical simulations of structures formation (for a complete description of the numerical simulation code, see Teyssier 2002). In order to simulate a merging cluster in a realistic cosmological environment, a low resolution ($N_{\text{part}} = 128^3$) large scale ($L_{\text{box}} = 400 h^{-1}$ Mpc)

Λ -CDM simulation was first performed. More than 100 large clusters of galaxies formed in this run. One cluster undergoing a violent merger was selected. Then, the exact same universe was re-simulated, adding many more particles and mesh points in the region surrounding the cluster. The same original box ($400 h^{-1}$ Mpc aside) is now described with 400 000 particles and 4 times more mesh points within the cluster virial radius. The mesh spacing in the cluster core has gone down to $6 h^{-1}$ kpc, corresponding to a formal resolution of 65536^3 . This “zoom” simulation lasted more than 100 CPU h on the Fujitsu VPP 5000 at CEA Grenoble. For this work, the cluster is always at the same comoving distance of $z = 0.1$. Just the time has increased from $z \approx 0.13$, $z \approx 0.09$, up to $z \approx 0.0$. This allows us to keep the same scale for all the images which is a square of 26.5 arcmin aside, i.e. a full EPIC-XMM Newton field of view. The projected gas emissivity and temperature of the merging cluster at each time are shown in Fig 9.

The map of averaged gas temperatures corresponding to the early stage of merging ($z \approx 0.13$) shows hot filaments and substructures ($T \approx 4$ keV) superimposed on a temperature background at 0.5 keV. The processes of gas heating and shocking are balanced by gas relaxation within each subgroup at $z \approx 0.09$, leading to two isothermal subgroups at respective temperature of 2 and 4 keV, where the hottest subgroup is the most massive. Eventually both subgroups have merged leading to a hot single one at $z = 0$. This merging cluster begins to reach its equilibrium temperature of 4 keV, with hot shocked regions superimposed from 5 to 6 keV. The temperature discontinuity between the central regions of the cluster and the background has increased during the merging process, since the central regions of the cluster have been heated whereas the background temperature remained constant.

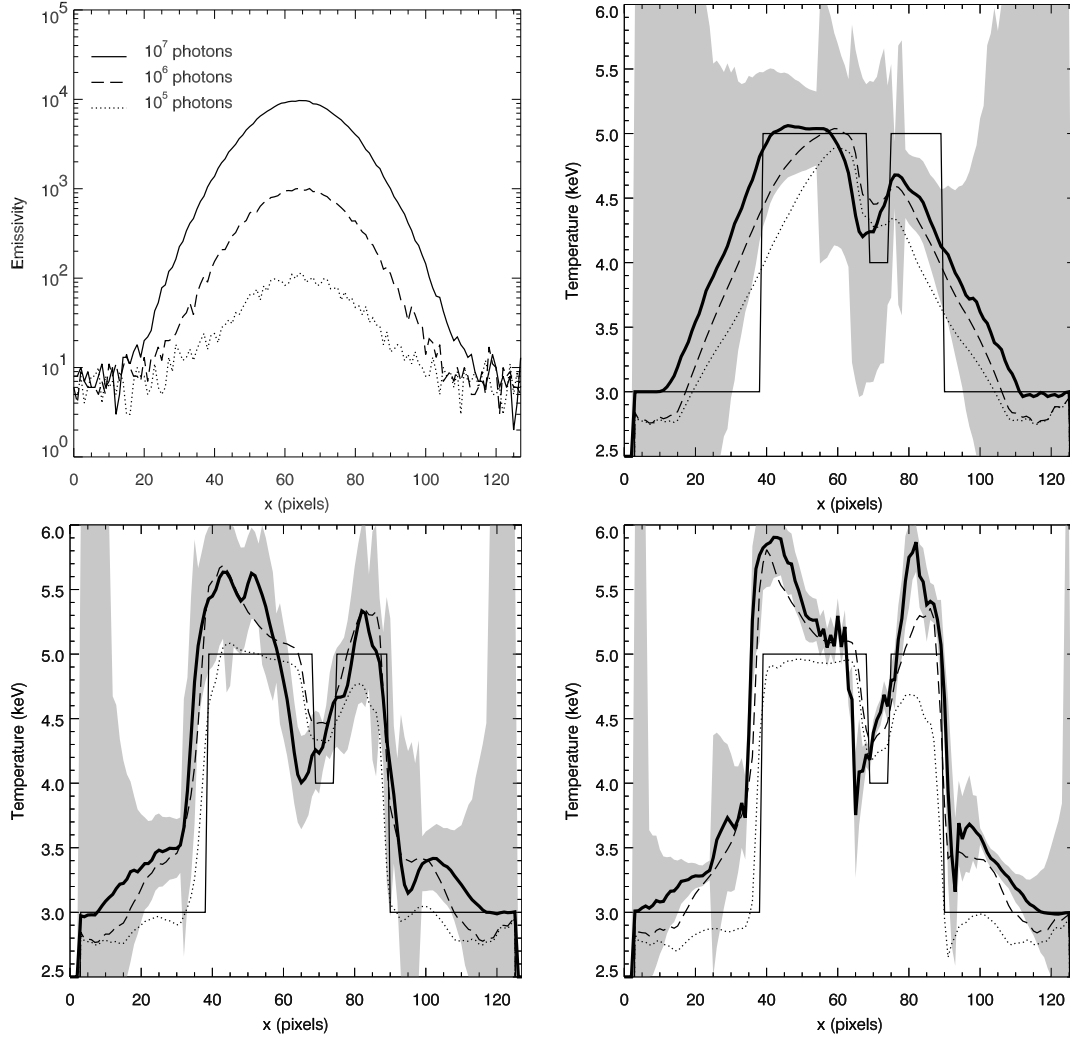


Fig. 8. Temperature profiles across the central region of simulated EPIC-XMM MOS1 observations, – temperatures coded in keV –. *Top-left:* emissivity profiles of the three emission models corresponding to global emissivities of 10^5 , 10^6 and 10^7 photons. *Top-right:* temperature profile with error-bars for a global emissivity value of 10^5 photons. *Bottom left and right:* temperature profiles with error-bars for respective global emissivity values of 10^6 and 10^7 photons. Bold line: temperature cuts, with grey shaded 3σ significance region superimposed. Thin line: temperature model. Dotted line: temperature model with the same wavelet thresholding as the temperature maps. Dashed line: temperature model with the same wavelet thresholding as the temperature maps and emissivity weighted mean temperature per bin.

5.2. Simulation of an EPIC-XMM MOS1 observation

The 3-D temperature $T(k, l, m)$ at pixel k , l , m , and emissivity $N_C(k, l, m)$ distributions of the intra-cluster plasma being known, an EPIC-XMM MOS1 observation can be simulated. To do so, the emitted photons are randomly drawn among different temperatures from the emission model of Poisson parameter $T(k, l, m)$ and have to follow the optical path from the source to the EPIC-XMM MOS1 focal plane. The different

distortions of the signal that would occur for a real cluster at $z = 0.1$ are taken into account: redshift, galactic absorption and instrumental response including the PSF, vignetting and energy response of the CCD detectors. The cluster and particle background emissivities correspond to standard emissivities for a real 5 keV cluster at $z = 0.1$ observed during 60 ks, yielding a background emissivity of $N_p = 0.67$. The CXB has not been taken into account. The simulated EPIC-XMM count maps with restored temperature maps are shown in Fig. 10.

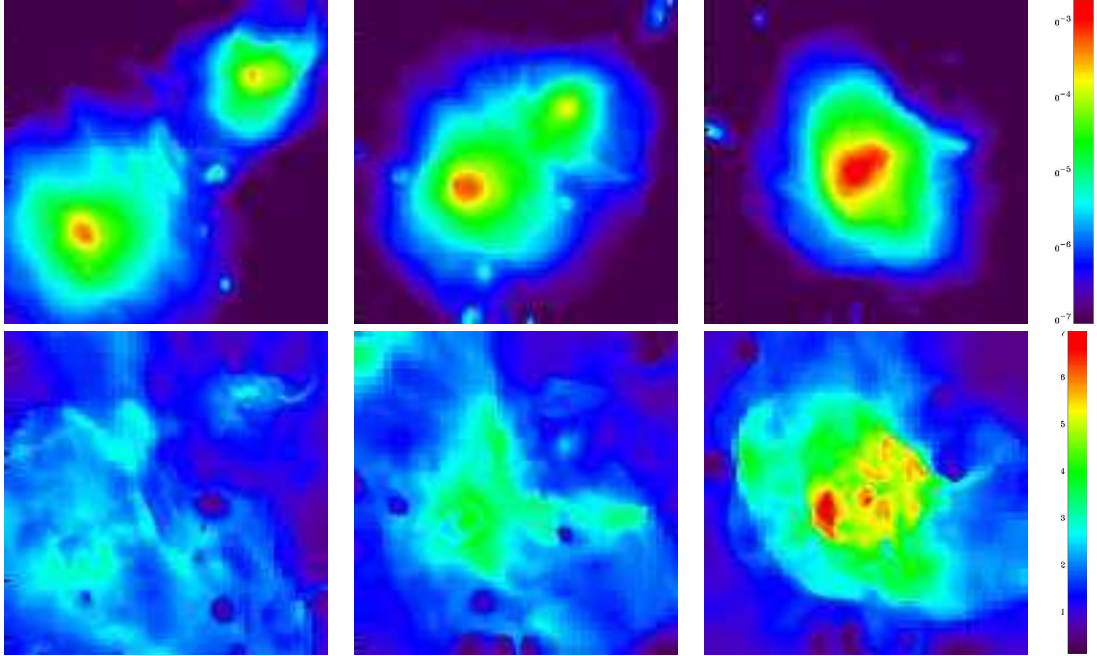


Fig. 9. Simulated merging cluster at $z \simeq 0.13$, $z \simeq 0.09$ and $z \simeq 0.0$ from left to right. *Top*: emissivity maps, with emissivity color code in $\text{ergs s}^{-1} \text{cm}^{-2}$. *Bottom*: maps of the mean gas temperature averaged along the line of sight, with temperature colors coded in keV. The width of the field of view is 26.5 arcmin.

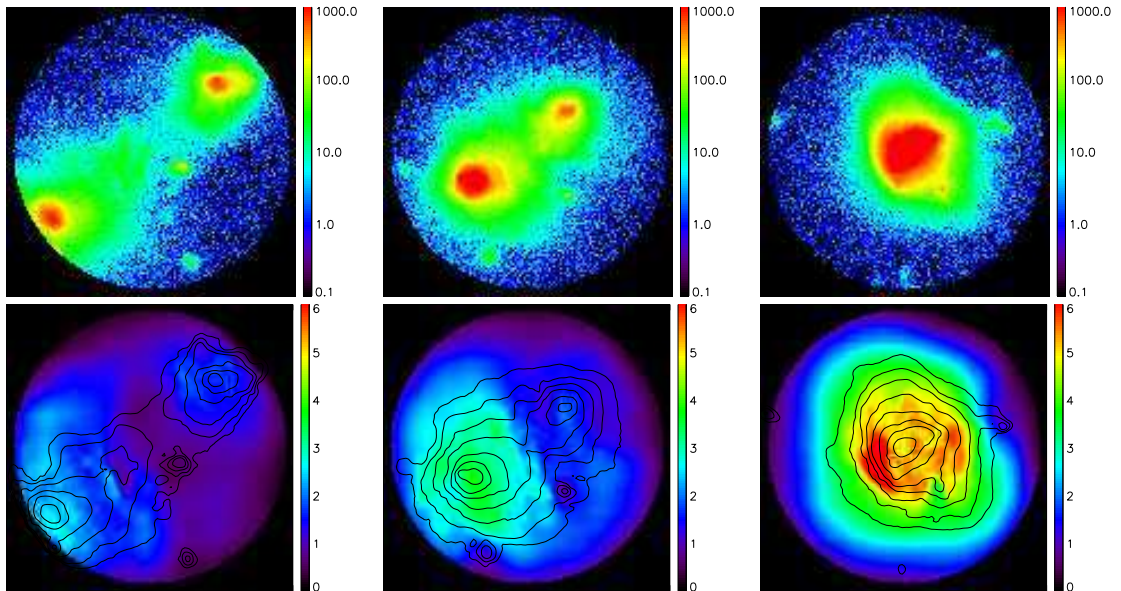


Fig. 10. Simulated merging cluster at $z \simeq 0.13$, $z \simeq 0.09$ and $z \simeq 0.0$ respectively. *Top*: simulated EPIC-XMM MOS1 photons counts maps, with emissivity coded in counts. *Bottom*: temperature maps with logarithmic emissivity contours superimposed, the temperatures are coded in keV.

5.3. Temperature restoration

The temperature maps of Fig. 10 have been restored with a 1σ significant threshold. This simulated observation is closer to the observation of a real cluster than the models of the previous section. Indeed the instrumental components are supposed to be similar, but the source emission is no longer a single component model due to projection effects. We consequently must bear in mind that the spectral fitting of each resolution element is not as accurate as described in Sect. 4.1. Let us nevertheless compare the restored temperature maps of Fig. 10 with the maps of averaged gas temperatures of Fig. 9.

The counts vary by several orders of magnitude between high and low density regions (cf. count maps of Fig. 10). As expected from a multiscale algorithm, the resolution of the temperature maps is better in regions with high emissivity and statistics, than in regions with poor statistics. For instance, the closer to the core of the subgroups, the better resolved are the details of the hot substructures of the early stage of merger ($z \simeq 0.13$). Likewise the hot shocked regions in the central region of the cluster at $z = 0$ are well restored despite the light smoothing applied, whereas both subgroups at $z \simeq 0.09$ and the single cluster at $z = 0$ have been detected at large scale as isothermal structures because of low statistics in the external regions.

These results on a simulated merging cluster at $z = 0.1$ allow us to conclude that our multiscale algorithm applied to real EPIC-XMM Newton observation of merging clusters should enable us to detect the expected temperature variations if they exist. Nevertheless the main building blocks of our algorithm, i.e. the spectral fitting and the multiscale temperature restoration, may still be improved. The spectral fitting uses a single component temperature model and a spectral modelling of projection effects with a multicomponent emission model might improve the detection of temperature variations. Also, the multiscale algorithm relies on a Haar wavelet which matches the square sampling of the data but leads to square patterns on the restored image, despite the regularization. Future improvements of the algorithm will include the use of other wavelets in order to better match the shape of the sources.

6. Conclusion

To take advantage of the spatial resolution of the new X-ray spectro-imagers, a new spectro-imaging algorithm has been introduced. This algorithm has been especially designed for the EPIC-XMM Newton cameras in order to restore the gas temperature variations within clusters of galaxies. The algorithm consists of two parts. First, an emission model of the source is fitted around each pixel of the detector surface, using regions of different areas, leading to a set of temperatures with error-bars at different scales. Then, the temperature variations are coded and thresholded in the wavelet space according to a significance criterion, and the wavelet transform of the optimal temperature map is obtained by regularizing the non-significant wavelet coefficients.

For the spectral fitting step, an emission model is fitted using a maximum likelihood method. The source emission

spectrum, the astrophysical and instrumental backgrounds, but also the instrumental response and especially the vignetting factor are taken into account within the emission model. As far as the vignetting is concerned, this approach is different from the usual solution, where the vignetting is taken into account by modifying the data instead of the emission model, e.g. the “weight method” described in Arnaud et al. (2001). The spectral fitting with full instrumental response modelling is more accurate in regions with very poor statistics and enables us therefore to reduce significantly the error bars in external regions of the FOV where the vignetting factor is important. However a temperature bias has been revealed in cases of poor statistics. This bias is enhanced in the case of high emissivity of the background with regard to the source emissivity, due to degeneracy effects in the shape of the emission spectra of each component. This result has to be kept in mind for sources like compact groups with low emissivities with regard to the background emissivity in the central region of the focal plane.

Our multiscale algorithm enables us to adapt the spatial resolution to the local statistics. In our test on a numerical simulation of a merging cluster at $z = 0.1$, the multiscale algorithm has enabled us to reveal not only faint temperature variations in the core of the merging subgroups, but also a large-scale temperature decrease in their external regions. The multiscale approach provides much more information than other methods using square or ring kernels. Nevertheless the results have still to be interpreted carefully. The final map has indeed been computed from a linear combination of spatially rebinned informations, and the knowledge of the area used for spectral fitting is not straightforward. Also, the chosen Haar wavelet transform is well suited to match the square spatial sampling but is not the best one since it implies block effects. Other wavelets that better match the temperature structures shape should be investigated.

Due to the spatial variation of the instrumental response, the different temperatures within different spatial bins are first computed from spectral fittings using different emission models. Thus, the wavelet transform of the expected temperature map is not computed linearly from the data. But a real wavelet transform is a linear process. This wavelet transform of the searched temperature map is obtained iteratively by regularizing the non-significant wavelet coefficients, which is questionable due to the non-linearity of the process. Yet this process has been tested on simulations and seems to perform well. The good behavior of the algorithm is due to the low spatial variations of the emission model, and especially to the low relative spatial variations of temperatures within the intra-cluster plasma. Consequently such an algorithm is well suited for restoring spatial variations of temperature within extended sources like clusters of galaxies but would not work for sources with high local variations of the emission model such as supernovae remnants.

The current spectro-imaging algorithm will be applied to a sample of merging clusters at different phases. Future developments are considered. The spectral fitting process could include spatial variations of the emission model in order to take into account projection effects. Moreover variations of other parameters such as the gas metallicity could be measured.

Appendix A: The maximum likelihood pursuit: A modified conjugate gradient method

The model fitting algorithm is based on the pursuit of the maximum of the likelihood function $L(r, T)$ – cf. Eq. (5) – over the 2-D space with parameters r and T . To do so, a 2-D minimization algorithm is required. The conjugate gradient method was chosen for its fast convergence, but adapted to a bounded parameter space with possibility of multiple maxima of the likelihood function.

Let us recall the general method of minimizationalgorithms. Considering a n -D function $S(X)$. The function $S(X)$ and its gradient can be developed over the initial position X_0 :

$$S(X) = S(X_0) + [X - X_0]^T \nabla_0 S(X_0) + \frac{1}{2} [X - X_0]^T H_0 [X - X_0], \quad (A.1)$$

$$\nabla S(X) = \nabla_0 S(X_0) + H_0 [X - X_0], \quad (A.2)$$

where H and ∇ are the respective Hessian and gradient operators and H_0 and ∇_0 their value at position X_0 .

The minimization is obtained for $\nabla S(X) = 0$. Since $S(X)$ is usually non-linear S has to be developed iteratively over successive positions X_i . Minimization are usually performed over 1-D cuts of $S(X)$. The successive directions of the cuts depending on the algorithms. In case of the conjugate gradient algorithm, a given position X_i corresponds to a minima in direction P_i . Then the next direction P_{i+1} is chosen to be a *conjugate* direction from P_i . Contrary to the simple gradient algorithm where P_{i+1} is normal to P_i , here P_{i+1} is chosen so that ∇S stays as far as possible normal to direction P_i , i.e. $P_i^T \delta[\nabla S(X)] = 0$. This condition prevents successive explorations in the same direction. By differentiation of Eq. (A.2) and by taking into account that $P_{i+1}^T \delta X = 0$ at this iteration, we get:

$$\delta[\nabla S(X)] = H_0 \delta X, \quad (A.3)$$

$$P_i^T H_0 P_{i+1} = 0. \quad (A.4)$$

Let us now describe the adaptation of this standard algorithm to our case with a 2-D bounded parameter space (cf. Fig. A.1) with possibility of multiple maxima. At iteration i , starting from X_{i-1} , a maximum of $L(r, T)$, X_i , is fetched along the conjugate direction from P_{i-1} : P_i . We are looking for the maximum of the 1-D cut of $L(r, T)$: $L_1(x)$ with origin $x = 0$ at point X_{i-1} . To do so we define a bounded exploration support $[x_{\min}, x_{\max}]$ of $L_1(x)$.

In any case, both bounds x_{\min} , x_{\max} must be included within the bounded parameter space, so that the exploration support $[x_{\min}, x_{\max}]$ must be included within the $[\tilde{x}_{\min}, \tilde{x}_{\max}]$ interval where \tilde{x}_{\min} and \tilde{x}_{\max} are the parameter space bounds along the direction P_{i-1} (cf. Fig. A.1). In case of good statistics the likelihood function $L(r, T)$ has only one maximum within the parameter space and the conditions $x_{\min} = \tilde{x}_{\min}$, $x_{\max} = \tilde{x}_{\max}$ are enough. But in the case of bad statistics, some multiple maxima events can appear, so that a choice has to be made among the possible solutions. We observed in simulations of observations that the good one is generally the closest to the center of the parameter space. Indeed two kinds of false maxima

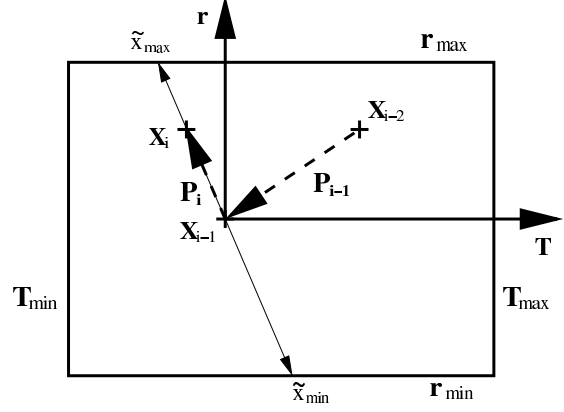


Fig. A.1. The bounded parameter space. At iteration i the minimum is obtained from X_{i-1} over direction P_i along the bounded segment $[\tilde{x}_{\min}, \tilde{x}_{\max}]$.

can appear on the T -axis. The first ones concern the low temperatures ($T < 1$ keV) in case of degenerescence due to the similar shape of the emission-lines complex and CXB emission in the soft band. The second ones concern the high temperatures ($T > 10$ keV) in case of degenerescence between the bremsstrahlung part of the cluster emission and CXB emission in the hard band. Consequently, the closest solution to the center of the parameter space, – where the algorithm starts its pursuit –, is chosen in the case of multiple maxima. To do so, a new constraint is added to the exploration support $[x_{\min}, x_{\max}]$ at iteration i : the right solution must be the closest to the starting point $x = 0$, so that x_{\min} and x_{\max} are randomly drawn in the respective $[\tilde{x}_{\min}, 0]$, and $[0, \tilde{x}_{\max}]$ intervals up to get values where $\frac{dL_1(x_{\min})}{dx} > 0$ and $\frac{dL_1(x_{\max})}{dx} < 0$. In case of bad statistics where no support $[x_{\min}, x_{\max}]$ can match the previous condition, the undefined value is returned for both parameters.

This local pursuit of function $L_1(x)$ is performed at each iteration of the standard conjugate gradient algorithm. In case of good statistics where a maximum is found, the convergence criterion $\|\nabla L(r, T)\| < \epsilon$ is obtained in a few iterations with $\epsilon = 1 \times e^{-6}$.

Appendix B: The 2-D Haar wavelet transform

The wavelet transform of a 2-D signal $I(k, l)$ is based on a wavelet function which can be dilated and translated. The 2-D wavelet leads to functions with compact support depending on three parameters: the “scale” or dilation factor a and both space coordinates k, l . In our case, for geometrical and computational reasons, the chosen wavelet is the Haar wavelet (cf. Fig. B.1).

Since the $I(k, l)$ distribution is sampled in the direct space a discrete wavelet transform $W(a, k, l)$ can be associated to $I(k, l)$ while preserving all the signal information. To do so the scale axis can be sampled within a dyadic scheme so that $a = 2^j$ with $j = 0, \dots, N$. The wavelet transform can be redundant within the space axes, such as in the so called “a trous”

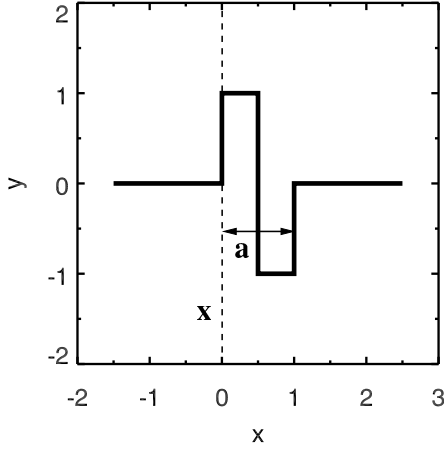


Fig. B.1. The 1-D Haar wavelet and its characteristic parameters scale (a) and location (x).

algorithm (Holschneider 1989), or non-redundant, such as in the Mallat multi-resolution algorithm (Mallat 1989).

On the one hand, the computation of the wavelet transform $W(a, k, l)$ can be performed from the 2-D signal $I(k, l)$. Given a scale a , “smoothed planes” $A(a, k, l)$ and wavelet coefficients $W(a, k, l)$ can be defined, both being computed from the smoothed plane $A(\frac{a}{2}, k, l)$ by applying low-pass $h(\frac{a}{2})$ and high-pass $g(\frac{a}{2})$ analysis filters, respectively. The smoothed plane $A(a, k, l)$ corresponds to the smoothing of $I(k, l)$ with a kernel of width a , whereas the wavelet coefficients $W(a, k, l)$ corresponds to the information lost between $I(k, l)$ and $A(a, k, l)$. Starting at scale $a = 1$, the first smoothed plane $A(1, k, l) \equiv I(k, l)$ can be defined, and the coefficients $W(a, k, l)$ and planes $A(a, k, l)$ can be computed iteratively from scale to scale.

The wavelet transform using the anisotropic 2-D Haar wavelet is composed of three kinds of coefficients, the horizontal, vertical and diagonal coefficients, $W_h(a, k, l)$, $W_v(a, k, l)$, $W_d(a, k, l)$, respectively. According to the redundant shift invariant wavelet transform (Coifman & Donoho 1995), they are computed by applying the corresponding, horizontal, vertical and diagonal analysis filters, of Fig. B.2, following the analysis equations:

$$A(a, k, l) = H\left(\frac{a}{2}\right)A\left(\frac{a}{2}, k, l\right), \quad (\text{B.1})$$

$$W_h(a, k, l) = G_h\left(\frac{a}{2}\right)A\left(\frac{a}{2}, k, l\right), \quad (\text{B.2})$$

$$W_v(a, k, l) = G_v\left(\frac{a}{2}\right)A\left(\frac{a}{2}, k, l\right), \quad (\text{B.3})$$

$$W_d(a, k, l) = G_d\left(\frac{a}{2}\right)A\left(\frac{a}{2}, k, l\right). \quad (\text{B.4})$$

On the other hand the computation of the 2-D signal $I(k, l)$ can be performed from its wavelet transform $W(a, k, l)$. According to the biorthogonal wavelet formalism, the so-called wavelet reconstruction is performed by applying at higher and higher resolution the respective low-pass $\tilde{h}(\frac{a}{2})$ and high-pass $\tilde{g}(\frac{a}{2})$

$$\begin{array}{c} \left[\begin{array}{cccccc} \frac{1}{2} & 0 & \dots & 0 & \frac{1}{2} \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ \frac{1}{2} & 0 & \dots & 0 & \frac{1}{2} \end{array} \right]_{2^{j+1}} \quad \left[\begin{array}{cccccc} \frac{1}{2} & 0 & \dots & 0 & \frac{1}{2} \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ -\frac{1}{2} & 0 & \dots & 0 & -\frac{1}{2} \end{array} \right]_{2^{j+1}} \\ \hline h(2^j) \quad \quad \quad g_h(2^j) \\ \left[\begin{array}{cccccc} \frac{1}{2} & 0 & \dots & 0 & -\frac{1}{2} \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ \frac{1}{2} & 0 & \dots & 0 & -\frac{1}{2} \end{array} \right]_{2^{j+1}} \quad \left[\begin{array}{cccccc} \frac{1}{2} & 0 & \dots & 0 & -\frac{1}{2} \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ -\frac{1}{2} & 0 & \dots & 0 & -\frac{1}{2} \end{array} \right]_{2^{j+1}} \\ \hline g_v(2^j) \quad \quad \quad g_d(2^j) \end{array}$$

Fig. B.2. The two dimensional analysis filters at scale 2^j . The low-pass $h(2^j)$ and the three high-pass $g_{h,v,d}(2^j)$ filters associated to the horizontal, vertical and diagonal details respectively.

$$\begin{array}{c} \left[\begin{array}{cccccc} \frac{1}{2} & 0 & \dots & 0 & \frac{1}{2} \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ \frac{1}{2} & 0 & \dots & 0 & \frac{1}{2} \end{array} \right]_{2^{j+1}} \quad \left[\begin{array}{cccccc} -\frac{1}{2} & 0 & \dots & 0 & -\frac{1}{2} \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ \frac{1}{2} & 0 & \dots & 0 & \frac{1}{2} \end{array} \right]_{2^{j+1}} \\ \hline \tilde{h}(2^j) \quad \quad \quad \tilde{g}_h(2^j) \\ \left[\begin{array}{cccccc} -\frac{1}{2} & 0 & \dots & 0 & \frac{1}{2} \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ -\frac{1}{2} & 0 & \dots & 0 & \frac{1}{2} \end{array} \right]_{2^{j+1}} \quad \left[\begin{array}{cccccc} \frac{1}{2} & 0 & \dots & 0 & -\frac{1}{2} \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ -\frac{1}{2} & 0 & \dots & 0 & \frac{1}{2} \end{array} \right]_{2^{j+1}} \\ \hline \tilde{g}_v(2^j) \quad \quad \quad \tilde{g}_d(2^j) \end{array}$$

Fig. B.3. The two dimensional synthesis filters at scale 2^j . The low-pass $\tilde{h}(2^j)$ and the three high-pass $\tilde{g}_{h,v,d}(2^j)$ filters associated to the horizontal, vertical and diagonal details respectively.

synthesis filters to the wavelet coefficients $W(a, k, l)$ and smoothed planes $A(a, k, l)$. The filtered wavelet and smoothed planes are added, and the right renormalization leads to a new smoothed plane $A(\frac{a}{2}, k, l)$, and so on at higher and higher resolution. The last smoothed plane at the highest resolution constitutes the 2-D signal $I(k, l)$.

In the case of the 2-D Haar wavelet transform, the smoothed plane $A(\frac{a}{2}, k, l)$ at scale $\frac{a}{2}$, is computed from a linear combination of the smoothed plane $A(a, k, l)$ and the wavelet coefficients $W_{h,v,d}(a, k, l)$, following the synthesis equation:

$$A\left(\frac{a}{2}, k, l\right) = \left[\tilde{H}A(a, k, l) + \tilde{G}_h W_h(a, k, l) + \tilde{G}_v W_v(a, k, l) + \tilde{G}_d W_d(a, k, l) \right]. \quad (\text{B.5})$$

References

- Arnaud, M., Neumann, D. M., Aghanim, N., et al. 2001, *A&A*, 365, L80
- Bobichon, Y., Bijaoui, A. 1997, *Exp. Astron.*, 7, 239
- Coifman, R. R., & Donoho, D. 1995, Translation invariant denoising, Technical report 475, Dept. of Statistics, Stanford University
- Colberg, J. M., White, S. D. M., Yoshida, N., et al. 2000, *MNRAS*, 319, 209
- Evrard, A. E. 1990, *ApJ*, 363, 349
- Evrard, A. E., MacFarland, T. J., Couchman, H. M. P., et al. 2002, *ApJ*, 573, 7
- Sanders, J. S., & Fabian, A. C. 2001, *MNRAS*, 325, 178
- Holschneider, M., Kronland-Martinet, R., Morlet, J., & Tchamitchian, P. 1989, A real-time algorithm for signal analysis with help of the wavelet transform, *Wavelets: Time frequency methods and phase-space*, 286
- Haar, A. 1910, Zur theorie der orthogonalen Funktionensysteme, *Mathematische Annalen*, 69, 331
- Jammal, G., & Bijaoui, A. 1999, A multiresolution image restoration method for photon imaging systems. *Proc. IEEE Intern. Conf. Image Processing*, 4, 3401
- Katz, N., & White, S. D. M. 1993, *ApJ*, 412, 455
- Kendall, M. G., Stuart, A. 1973, *The advanced theory of statistics*, T.2 (Londres: Griffin)
- Lumb, D. H., Warwick, R. S., Page, M., & De Luca, A. 2002, *A&A*, 389, 93
- Mallat S. 1989, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. on pattern analysis and machine intelligence*, 11, 674
- Mewe, R. 1972, *SpH*, 22, 459
- Mewe, R. 1975, *SpH*, 44, 383
- Mewe, R., Gronenschild, E. H. B. M., & van den Oord, G. H. J. 1985, *A&AS*, 62, 197
- Mewe, R., Lemen, J. R., & van den Oord, G. H. J. 1986, *A&AS*, 65, 511
- Ricker, P. M., & Sarazin, C. 2001, *ApJ*, 561, 621
- Ritchie, W. R., & Thomas, P. A. 2002, *MNRAS*, 329, 675
- Roettiger, K., Burns, J. O., & Loken, C. 1996, *ApJ*, 473, 651
- Sarazin, C. L. 1986, *Rev. Mod. Phys.*, 58, 1
- Teyssier, R. 2002, *A&A*, 385, 337
- Tikhonov, A. N. 1963, Regularization of incorrectly posed problems, *Sov. Math. Dokl.*, 4, 1624
- Van Cittert P. H. 1931, *Z. Phys.*, 69, 298

4.5 Conclusion.

Nous avons construit un algorithme de spectro-imagerie multi-échelles adapté à la restauration des variations de température et d'entropie du milieu intra-amas. L'estimateur local utilise la méthode du maximum de vraisemblance et permet ainsi de bénéficier de la résolution spatiale offerte par les spectro-imageurs EPIC. Nous avons vérifié que cet estimateur se comporte comme un estimateur optimal atteignant la borne de Rao-Cramer dans les conditions courantes d'utilisation. Il peut toutefois être fortement biaisé dans les régions dominées par les composantes de fond et son utilisation dans de telles circonstances doit être validée au cas par cas. L'algorithme de reconstruction multi-échelles est basé sur le seuillage d'une transformée en ondelettes de Haar redondante, qui est adaptée à la géométrie des grilles dans lesquelles on procède à l'estimation locale. Il n'existe pas au départ de relation linéaire entre les coefficients à différentes échelles, ce qui nécessite d'introduire un processus de linéarisation. Du fait de cette non linéarité initiale, la pondération des estimations de température par la distribution d'émissivité peut conduire à des incohérences entre les estimations à différentes échelles si les gradients de température sont trop forts, puis à une mauvaise estimation du paramètre final. En outre, les effets de projection ne sont pas pris en compte dans l'estimation. Malgré ces réserves, la simulation d'une observation d'amas en cours de fusion dans des conditions proches des conditions réelles a permis de montrer que cet algorithme pouvait mettre en évidence de manière satisfaisante des structures caractéristiques des processus de fusion dans le milieu intra-amas.

Si cet algorithme fournit déjà des résultats satisfaisants dans sa version actuelle, nous pouvons néanmoins en suggérer des améliorations en vue de développements futurs. Dans l'algorithme d'imagerie, nous avons utilisé l'ondelette de Haar parce que sa géométrie est adaptée à l'estimation dans des grilles uniformes. Une autre fonction d'échelle conduisant à une ondelette moins discontinue pourrait toutefois diminuer quelques artefacts de reconstruction. Une déconvolution de la PSF d'XMM pourrait aussi être envisagée. En outre, nous avons discuté du problème posé par les fortes variations d'émissivité des amas de galaxies, à la fois sur les erreurs d'estimation et sur la non-prise en compte d'une grande partie de l'espace spatio-fréquentiel. Ce problème pourrait être résolu en adaptant l'échantillonnage à l'émissivité de l'amas. Toutefois, on introduirait alors un a priori sur la géométrie. Dans l'algorithme de spectroscopie, nous avons supposé un modèle d'émission isotherme à une composante thermique. Nous pourrions prendre en compte les effets de projection en modélisant une superposition de structures isothermes de géométrie simple, par exemple deux amas sphériques isothermes en cours de fusion. Par ailleurs, un estimateur de métallicité devra être implémenté. Enfin, le temps de calcul de l'estimateur pourrait être singulièrement réduit en n'estimant qu'un seul paramètre ou en projetant la bibliothèque de spectres dans une base de composantes principales.

Troisième partie

Applications

Chapitre 1

Restauration d'images

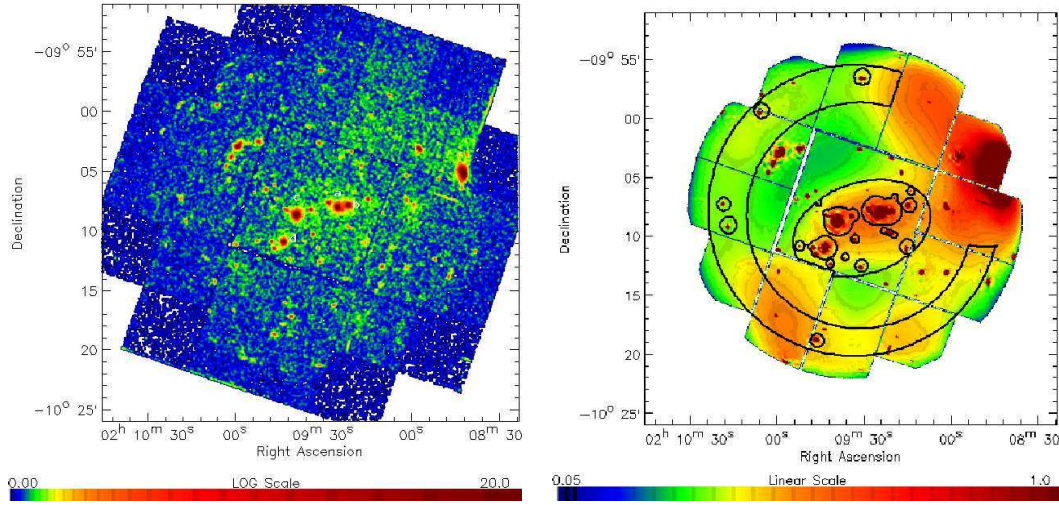


FIG. 1.1 – *Emission étendue du groupe compact HCG 16. Gauche: Distribution poissonnienne de photons dans la bande 0.2 - 7 keV. Droite: Distribution d'émissivité du plasma restaurée par l'algorithme de Jammal & Bijaoui.*

1.1 HCG 16 / Emission étendue dans un groupe compact.

Les groupes de galaxies sont des systèmes moins massifs que les amas, constitués de quelques dizaines de galaxies au plus. Ce sont des structures de taille intermédiaire entre les galaxies et les amas, présentant des propriétés variées suivant leur masse et leur état d'évolution. Si les groupes pauvres en galaxies spirales sont généralement considérés comme gravitationnellement liés, cette propriété est débattue dans les groupes de spirales. Comme dans les amas, c'est souvent la détection d'une émission X diffuse du plasma entre les galaxies qui prouve que les groupes sont gravitationnellement liés.

L'étude des groupes est intéressante dans le contexte de la formation des structures puisqu'ils constituent l'étape intermédiaire entre les galaxies et les amas. En outre, le plasma intergalactique présente des propriétés particulières dans les groupes, favorisant l'étude des phénomènes non gravitationnels. Comme la masse des groupes est plus faible que celle des amas, la température d'équilibre du plasma piégé dans le potentiel gravitationnel est plus basse. En conséquence, le plasma intergalactique n'est pas complètement ionisé dans les groupes et son spectre d'émission se caractérise par une contribution relative des raies par rapport au rayonnement de freinage plus importante que dans les amas. Par ailleurs la dispersion de vitesse du gaz étant plus faible que dans les amas, la dynamique du gaz est plus sensible aux interactions avec les galaxies. L'émissivité du plasma intra-groupe a été étudiée dans une trentaine de groupes brillants en X observés à l'aide du satellite ROSAT par Heldson & Ponman (2000). Les auteurs ont montré que le milieu intra-groupe est caractérisé par une distribution d'émissivité plus plate que le milieu intra-amas et une relation température-luminosité beaucoup plus pentue, $L_X \propto T^{4.9}$. Le potentiel gravitationnel des groupes étant moins profond que dans les amas, ces propriétés s'interprètent par une contribution plus importante des phénomènes non gravitationnels à l'augmentation d'énergie interne du plasma. En particulier, les vents galactiques sont supposés augmenter l'entropie du plasma avant l'effondrement gravitationnel.

Le groupe de galaxies HCG16 est un groupe compact constitué de sept galaxies spirales. Outre l'émissivité X des galaxies actives du groupe, une émission diffuse perturbée avait été détectée loin des galaxies avec des données ROSAT/PSPC par Ponman et al. (1999) et Dos Santos & Mamon (1999). Cet amas a constitué l'une des toutes premières cibles du satellite XMM-Newton.

L'émissivité des galaxies actives du groupe a été étudiée par Turner et al. (2001) et celle de la composante diffuse par Belsole et al. (2003). Dans cette dernière étude l'algorithme de Jammal & Bijaoui a été utilisé pour restaurer la distribution spatiale de l'émissivité du plasma à partir de la liste d'événements poissonnienne obtenue avec les caméras MOS1 et MOS2. L'émission des galaxies étant particulièrement intense par rapport à l'émission diffuse, – plus de trois ordres de grandeur d'écart –, l'application de l'algorithme générait des artefacts de seuillage dus à la contribution des galaxies à basse fréquence, dont l'amplitude était comparable à celle de l'émission étendue. Pour supprimer ces artefacts, la contribution des galaxies a été soustraite avant de procéder au lissage, puis réintroduite sur le résultat final. Une émission diffuse elliptique de gaz chaud entourant le sous-groupe a alors été mise en évidence (cf. fig. 1.1). Cette structure est beaucoup plus régulière que la distribution des galaxies ou que la structure détectée par Dos Santos & Mamon avec les données ROSAT. Ce résultat est dû à la meilleure sensibilité d'XMM qui a permis d'identifier puis d'extraire des sources ponctuelles, mais aussi à la technique de lissage employée. En effet Dos Santos & Mamon avaient utilisé un algorithme de lissage adaptatif implémenté avec un seuil de signification plus élevé que la variance de la distribution poissonnienne associée à l'émission étendue. La distribution elliptique du gaz mise ainsi en évidence a permis d'isoler spatialement la région centrale où l'émissivité du plasma domine le fond et d'en extraire le spectre d'émission. Une température de 0.49 keV et une métallicité de $0.07 Z_{\odot}$ ont été obtenues par ajustement. Le rapport luminosité-température est en accord avec la loi auto-similaire observée par Heldson & Ponman pour les groupes de galaxies (cf. fig. 6 de l'article de Belsole et al.).

Le fait que la structure étendue soit régulière et détectée à grande distance du centre des galaxies actives (135 kpc) tend à prouver que le groupe constitué exclusivement de spirales, est gravitationnellement lié, ce qui rouvre le débat sur l'état dynamique de ce type d'amas. Peut-être constitue-t-il un cas exceptionnel de groupe en train de se relaxer après une formation récente, dans lequel les galaxies de type tardif n'ont pas interagi longtemps avec le plasma.

Chapitre 2

Spectro-imagerie

2.1 Deux amas de galaxies bi-modaux en coalescence: Abell 3921 et Abell 1750

Les amas de galaxies Abell 3921 et Abell 1750 ont été observés durant le temps garanti XMM dans le cadre d'un programme d'observation d'amas en coalescence (P.I. J.-L. Sauvageot, CEA). Ils ont été sélectionnés à partir d'observations ROSAT pour leur morphologie bimodale et leur extension spatiale couvrant le champ angulaire de 30 minutes d'arc des caméras EPIC. Dans une configuration où les deux sous-groupes de chaque amas chutent l'un sur l'autre, on s'attendait en particulier à mettre en évidence une onde stationnaire de compression se développant à égale distance des coeurs de chaque sous-groupe, caractérisée par une structure en barre avec une élévation significative de la température du gaz. Ces amas sont parallèlement étudiés en optique dans le cadre d'un programme d'observation multi-longueurs d'onde (P.I. S. Maurogordato, OCA). La cartographie des distributions de température du plasma a clairement mis en évidence les ondes de compression attendues, mais dans le cas d'Abell 3921 il ne s'agit sans doute pas d'une configuration où les deux groupes chutent l'un sur l'autre. Dans le cas d'Abell 1750, on observe en outre des ondes de choc dans l'un des sous-groupes, probablement associées à des processus de coalescence antérieurs au processus actuel. Dans chaque cas, c'est une analyse conjointe des distributions de température et d'entropie du plasma à deux dimensions et des profils de densité et de température à une dimension de chaque sous-groupe qui a permis d'élaborer la dynamique du scénario probable de coalescence (Belsole et al., 2004).

2.1.1 A3921

Distribution de température du plasma

L'amas de galaxies Abell 3921 est constitué de deux sous-groupes en coalescence séparés de 1,2 Mpc ($h_o = 0.5$) et dont le rapport de masse est de l'ordre de $\frac{1}{2}$ à $\frac{1}{2.5}$ (Ferrari et al., 2004). La figure 2.1 est une superposition de la distribution de température et des isocontours d'émissivité du plasma intra-amas. Le détail de l'analyse multiéchelle qui a conduit à la carte de température est discuté dans la section 4.3.4 de la partie II.

La structure bimodale de l'amas apparaît clairement à grande échelle. On distingue un groupe principal de forme elliptique régulière et un groupe plus perturbé au Nord-Ouest¹, lui-même composé de deux sous-groupes séparés par une dépression. La morphologie générale de la distribution de gaz coïncide remarquablement avec la distribution de la densité numérique de galaxies (Ferrari et al., 2004), avec son groupe principal orienté Est-Ouest et un sous-groupe plus perturbé orienté Sud-Ouest/Nord-Est. Par ailleurs les trois pics de densité X correspondent aux trois galaxies les plus brillantes de l'amas.

La carte de température met en évidence une barre de compression entre les deux sous-groupes, caractérisée par une élévation significative de température. La température moyenne du gaz est de 5 keV dans le groupe principal, tandis qu'elle s'élève à 8 keV dans la barre de compression. Correspondant à un nombre de Mach de $M = 1.6$, la vitesse du gaz choqué de part et d'autre de l'onde de compression est de l'ordre de 840 km.s^{-1} (cf. equation (4.9)). On observe par ailleurs que la distribution de température dans la région centrale du groupe principal est très perturbée et que le profil radial de température de ce même groupe décroît fortement de 5 à 3.5 keV dans la partie Est opposée à l'onde de compression. La distribution perturbée de la température dans la région centrale et l'absence de courant de refroidissement montrent que le plasma n'est pas relaxé dans le groupe principal. Par ailleurs les simulations numériques semblent montrer que des profils en température décroissant fortement s'observent plutôt dans les amas en interaction. Une forte élévation de température ($kT > 8 \text{ keV}$) est enfin constatée dans la dépression séparant les deux sous-groupes au Nord-Ouest. Une source radio étant localisée dans la dépression, une interprétation possible serait que la dépression s'est développée autour de la source, le gaz étant éjecté radialement par la pression de radiation de la source. Son élévation de température pourrait

1. Les conventions utilisées pour les points cardinaux sont celles de la sphère céleste qui est usuellement vue de l'intérieur contrairement à la sphère terrestre. L'axe Nord-Sud reste le même, mais l'Est et l'Ouest sont inversés.

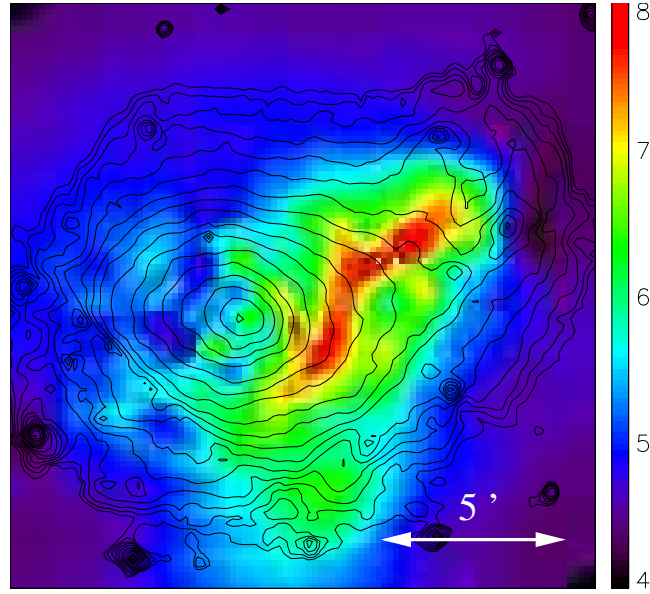


FIG. 2.1 – *Coalescence bimodale de deux sous-groupes dans l’amas de galaxies Abell 3921. Distribution de température (keV) et isocontours d’émissivité (échelle logarithmique). Pour un décalage spectral de $z = 0.094$ et avec $H_0 = 75 \text{ km.s}^{-1} \cdot \text{Mpc}^{-1}$, $\Omega_M = 0.3$, et $\Omega_\Lambda = 0.7$, 5 minutes d’arc correspondent à 489 kpc.*

alors s’expliquer par la compression de la coquille de gaz en expansion autour de la source, ou encore par une émissivité non thermique en partie due à la source elle-même. Si la contribution de la source n’est pas à exclure, il semble cependant que l’extension de la dépression et de la zone chaude soient trop importantes pour n’être imputables qu’à la seule source.

Scénario de coalescence.

Le scénario de coalescence de l’amas peut-être interprété en comparant les observables de spectro-imagerie X avec des simulations numériques d’une part et des observations optiques obtenues par Ferrari et al. (2004) d’autre part. Une première information délivrée par les observations optiques est que l’axe joignant les deux groupes en coalescence est quasiment dans le plan du ciel. Ce résultat s’observe qualitativement en constatant que la dispersion des composantes radiales de vitesse des galaxies de chaque groupe est gaussienne et ne présente aucune signature d’effets de coalescence, alors que l’état dynamique du milieu intra-amas montre que les groupes sont en interaction. Les deux groupes en coalescence étant alignés sur le plan du ciel, il est facile de comparer la distribution de température du plasma avec les prédictions de simulations numériques. On observe que l’onde de compression qui s’est développée entre les deux groupes n’est pas orthogonale à l’axe joignant leurs centres respectifs. Ce résultat exclut a priori l’hypothèse d’une collision avec un paramètre d’impact nul dans laquelle les deux groupes sont en train de chuter l’un sur l’autre. La comparaison de la distribution de température observée avec des simulations d’amas en coalescence bimodale (Ricker & Sarazin, 1995) montre que cette configuration peut par contre s’interpréter dans le cas d’une diffusion de paramètre d’impact non nul, les sous groupes étant déviés l’un par l’autre en générant une onde de compression du gaz en forme de “Z” à leur interface, alignée avec la direction de déviation. Le groupe Nord-Ouest aurait donc été diffusé par le groupe principal et s’en éloignerait maintenant en direction du Nord. Notons que sous cette hypothèse, la forte élévation de température observée au Nord du groupe Nord-Ouest pourrait constituer une onde de choc en avant du groupe. La torsion de l’onde de compression s’interpréterait alors par le fait que sa partie Ouest continue à se propager librement vers le Nord en avant du sous-groupe

Nord-Ouest alors que sa partie Est s’est arrêtée au contact du groupe principal. Une échelle de temps de la diffusion des deux groupes inférieure à 0.6 milliard d’années pourrait alors être estimée à partir de la vitesse de propagation de l’onde de choc, $v_c = 840 \text{ km.s}^{-1}$, dont la torsion se serait déroulée sur une distance de l’ordre de la longueur de la barre centrale du “Z”: 0.5 Mpc.

Ce scénario de coalescence est concordant avec l’analyse des observations optiques de Ferrari et al. (2004). En effet, les auteurs ont comparé les distributions de densité de galaxies et de gaz du groupe Nord-Ouest et montré que le pic de densité de galaxies est légèrement décalé au Nord-Est du pic de densité de gaz. On s’attend en effet à ce que la composante non collisionnelle de matière du groupe Nord-Ouest (matière noire + galaxies) soit en avance par rapport à la distribution de gaz dans le mouvement d’éloignement. En outre, les auteurs ont cherché les solutions cinématiques envisageables pour obtenir la distribution bimodale de vitesses de galaxies observée dans le cas d’un problème à deux corps. Ils ont montré que le scénario le plus probable est celui dans lequel les deux groupes alignés dans le plan du ciel sont en train de s’éloigner après s’être croisés il y a 0.25 milliard d’années, ce qui est cohérent avec les échelles de temps données par l’analyse de la dynamique du plasma intra-amas. Enfin ils ont montré que la région où se développe l’onde de compression est caractérisée par un excès significatif du nombre de galaxies à fort taux de formation d’étoiles, phénomène imputable à l’interaction des galaxies avec le plasma choqué sur des échelles de temps courtes (0.1 milliard d’années) correspondant aux échelles de temps du scénario global.

2.1.2 A1750

L’amas de galaxies Abell 1750 est un amas en coalescence bimodale, constitué de deux sous-groupes de masses voisines, déjà observé par les satellites Einstein (Forman, 1981), ROSAT et ASCA (Novicki, 1998, Donnelly, 2001). Donnelly et al. avaient à la fois analysé la dispersion de vitesses des galaxies de chaque sous-groupe à partir d’observations spectrographiques au Las Campanas Observatory et cartographié la distribution de température du plasma intra-amas à partir d’observations ROSAT et ASCA. Les auteurs avaient montré que les dispersions de vitesses des deux sous-groupes étaient caractéristiques de deux systèmes isolés, avec une différence de vitesse radiale de l’ordre de 1300 km.s^{-1} . Par ailleurs, une élévation de température du plasma de l’ordre de 30 % avait été identifiée à l’interface des deux sous-groupes, ainsi qu’une faible élévation de densité par rapport à une configuration où les deux sous-groupes seraient isolés. Les auteurs avaient conclu au développement d’une région de choc de compression en début de phase de coalescence. La sensibilité et la résolution spatiale du satellite XMM-Newton ont permis à Belsole et al. (2004) de confirmer et d’affiner ce scénario dans le cadre d’une analyse détaillée des distributions spatiales d’émissivité, de température et d’entropie du plasma intra-amas.

La figure 2.2 représente les distributions spatiales de température et d’entropie de l’amas, avec les isophotes d’émissivité superposées. Une carte d’erreur ainsi qu’une comparaison des résultats de notre algorithme de spectro-imagerie avec ceux obtenus en utilisant un algorithme d’échantillonnage adaptatif sont visibles à la page 8 de l’article de Belsole et al.. Les distributions de température et d’émissivité permettent globalement de distinguer trois régions principales:

- Le sous-groupe Nord, le moins massif et le plus froid, caractérisé par une distribution de température légèrement perturbée.
- Une région de compression chaude à l’interface des deux sous-groupes.
- Le sous-groupe central², caractérisé par une distribution de température fortement perturbée.

Les températures de chacune de ces régions ont été estimées à partir de listes d’événements extraites de la liste globale (cf. fig. 3 et 4 de l’article de Belsole et al.). Elles sont respectivement de 3.87 ± 0.1 , 2.84 ± 0.1 et $5.12 \pm 0.7 \text{ keV}$. La distribution de température du sous-groupe central est elle-même séparable en une région fortement perturbée à l’Est et une région relativement isotherme

2. Cette dénomination suit celle de Beers et al. (1991) qui ont identifié en optique un troisième sous-groupe moins massif au Sud des deux premiers, hors du champ de vue de cette observation.

à l'Ouest. Les auteurs ont établi des profils annulaires d'émissivité, de température et d'entropie du sous-groupe Nord et des régions Est et Ouest du sous-groupe central.

Le scénario de coalescence global est celui proposé par Donnelly et al., dans lequel un front de choc caractérisé par une légère élévation de température se forme à l'interface des deux sous-groupes au début d'une phase de coalescence. La température de la région choquée est de $kT = 5.12 \pm 0.7$ keV, correspondant à une élévation de température de 40 à 90 % par rapport à la température moyenne des régions de part et d'autre du choc ($kT = 3.10 \pm 0.14$ keV). Puisqu'il s'agit de deux groupes de masses sensiblement égales la région de choc est quasiment au repos, proche du centre de masse de la collision. L'application des conditions de saut d'Hugoniot-Rankine permet de déterminer le facteur de compression puis le nombre de Mach du choc, $M = 1.64$ (cf. équation (4.10)). La vitesse relative des deux groupes, déduite de l'équation (4.9) avec $v_c = 2 \times \Delta v_s$, est de 1400 km.s^{-1} . Cette valeur proche de l'écart moyen de dispersion de vitesses des galaxies de chaque sous-groupe suggère que l'axe de la collision est voisin de la ligne de visée. Dans cette configuration la distance de séparation entre les sous-groupes est inconnue. L'état d'avancement de la coalescence comparé aux simulations numériques de Ricker & Sarazin (2001) suggère néanmoins qu'elle est supérieure à la distance projetée de $675 h^{-1}$ kpc, -c'est-à-dire que les sous-groupes ne sont pas dans un plan orthogonal à la ligne de visée-, mais qu'elle est inférieure au rayon Viriel du sous-groupe Nord. Ces deux valeurs limites permettent d'estimer que les deux sous-groupes se croiseront dans 0.5 à 1.3 milliards d'années.

Les distributions et profils de température du sous-groupe Nord sont caractéristiques d'un amas relaxé éventuellement légèrement perturbé par l'interaction. Les profils de température et d'émissivité évoquent la possibilité d'un courant de refroidissement central. Les auteurs ont discuté cette hypothèse en évaluant l'éventuel temps de refroidissement et les possibilités d'une émissivité multiphase du plasma au coeur du sous-groupe. Il s'agirait le cas échéant d'un faible courant de matière, mais cette interprétation n'est pas définitive parce qu'elle ne prend pas en compte l'influence d'une structure au Nord du sous-groupe non visible dans le champ de vue de l'observation.

Le sous-groupe central, dont la distribution de température est fortement perturbée, n'est pas relaxé. Les auteurs avancent un faisceau d'indices suggérant qu'il a probablement subi des processus de coalescence violents antérieurs au processus en cours entre les deux sous-groupes principaux. Ces indices sont la mise en évidence d'une discontinuité de densité associée à la propagation d'une onde de choc, la variation d'orientation des isophotes de densité, l'absence de courant de refroidissement et l'excès d'entropie dans la région centrale. Cet excès a été mis en évidence en comparant l'allure du profil d'entropie du sous-groupe central avec celle des profils autosimilaires observés dans les amas relaxés (cf. figure 12 de l'article de Belsole et al.). Ce type d'excès est prévu dans les amas ayant subi des processus de coalescence par les simulations numériques de Ricker & Sarazin (2001) et Ritchie & Thomas (2002). Il est dû à la contribution de mouvements de convection qui ramènent le gaz choqué des régions externes vers les régions centrales et subsiste pendant des échelles de temps bien plus longues que les temps de traversée des amas. L'allure perturbée de la distribution d'entropie du sous-groupe central est en accord avec ce scénario.

Les auteurs ont ainsi pu mettre en élaborer un double scénario de coalescence, avec d'une part la collision des deux sous-groupes qui devraient se rencontrer dans environ un milliard d'années et d'autre part un sous-groupe central non relaxé résultant de processus de coalescence datant de un à deux milliards d'années. Ce scénario de collisions multiples est d'autant plus probable que l'environnement de l'amas de galaxies Abell 1750 est relativement dense. En effet, on observe à plus grande échelle que les trois sous-groupes constitutifs de l'amas sont alignés avec d'autres structures suggérant une configuration filamentaire, elle même membre d'un super-amas.

2.2 Ecarts à l'isothermalité dans un amas massif: Abell 2163

Situé à $z = 0.201$, l'amas de galaxies A2163 est un amas riche ($N_{Abell} = 119$ galaxies) et massif, par ailleurs très lumineux en X ($L_X = 6.10^4$ ergs) et exceptionnellement chaud. Sa morphologie X est régulière mais présente des asymétries prononcées avec une orientation ellipsoïdale variant avec le rayon et une compression des isophotes au Sud-Ouest.

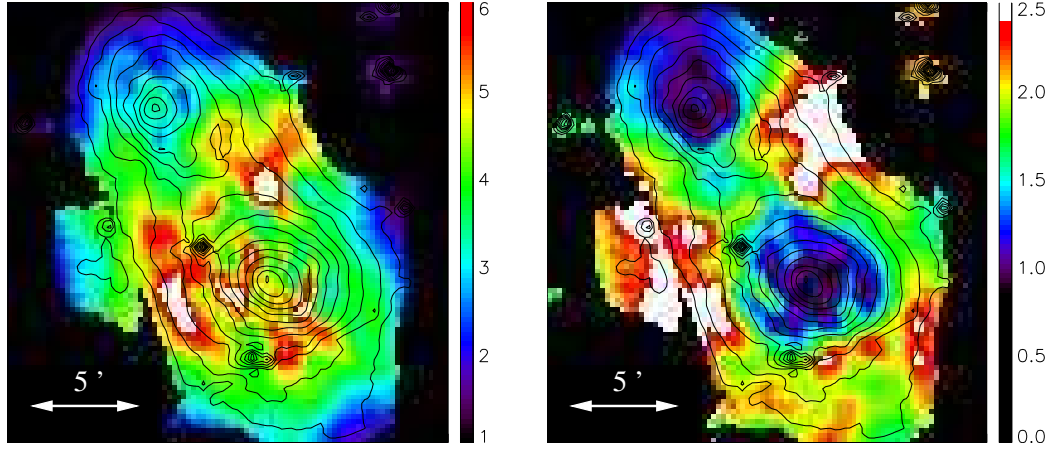


FIG. 2.2 – *Abell 1750. Distributions de température (Gauche) et d'entropie du plasma obtenues avec les caméras EPIC-MOS. (Belsole et al., 2004). Pour un décalage spectral de $z = 0.086$ et avec $H_0 = 75 \text{ km.s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$, et $\Omega_\Lambda = 0.7$, 5 minutes d'arc correspondent à 451 kpc.*

Elbaz et al. (1995) ont évalué sa température moyenne à $kT = 14.6 \text{ keV}$ à partir de données ROSAT-PSPC. Sous l'hypothèse d'équilibre hydrostatique, ils ont déduit une masse totale de $4.6 \times 10^{15} M_\odot$ et une fraction de baryons de 0.3. Ces valeurs sont exceptionnellement élevées, même pour un amas massif. A titre de comparaison, l'amas de Coma a une masse totale de $1.4 \times 10^{15} M_\odot$ et une température moyenne de $kT = 8 \text{ keV}$. Les auteurs ont discuté des implications cosmologiques qu'auraient ces valeurs si elles étaient fréquemment observées dans les amas massifs. En particulier, une fraction de baryons trop élevée est en désaccord avec le modèle de formation des structures dominé par la matière sombre froide. Mais ils ont aussi mentionné que cet amas présentait les propriétés d'un amas massif ayant récemment subi des collisions par des sous-groupes et n'étant pas encore relaxé. Ces propriétés sont les suivantes:

- la variation de l'orientation ellipsoïdale avec le rayon, similaire à celle des amas post-fusion non relaxés obtenus dans les simulations numériques (Schindler & Muller, 1993);
- la vitesse de dispersion élevée et la distribution perturbée des galaxies (Soucaill, 1995);
- le décalage entre la position du centroïde X et celle de la galaxie la plus brillante de l'amas;
- la présence d'un halo radio lumineux et étendu, caractéristique des amas massifs ayant subi des coalescences de sous-groupes récentes (Feretti, 2001);
- l'absence de courant de refroidissement central;
- la forte décroissance de la température X avec le rayon (Markevitch (1994));

Dans ces conditions, si l'amas A2163 reste très chaud et très massif, le caractère exceptionnel des propriétés observées est sans doute en partie dû au caractère non relaxé de l'amas. En effet, on a vu que les processus de fusion sont caractérisés par des phases de fortes élévation de la température moyenne du plasma avant la dissipation des ondes de choc (cf. Ricker & Sarazin, 2002, section 4.2.1). Par ailleurs, Markevitch et al. (1994) ont discuté le fait que la masse totale et la fraction de baryons déduites dans cet amas sous l'hypothèse d'équilibre hydrostatique pouvaient être surévaluées. En effet, avant la relaxation la distribution de vitesse du gaz reste perturbée et d'amplitude plus élevée qu'à l'équilibre, impliquant un excès de gaz à grand rayon. Par ailleurs le profil de température ne semble pas isotherme.

Nous avons cartographié la distribution de température de l'amas de galaxies A2163 à partir de données EPIC-MOS obtenues dans le cadre du temps garanti XMM (P.I. M. Arnaud, CEA). Cette distribution confirme à la fois la valeur élevée de la température moyenne ($T \simeq 14 \text{ keV}$) du gaz mais aussi la nature non relaxée de l'amas, en mettant en évidence des variations de température

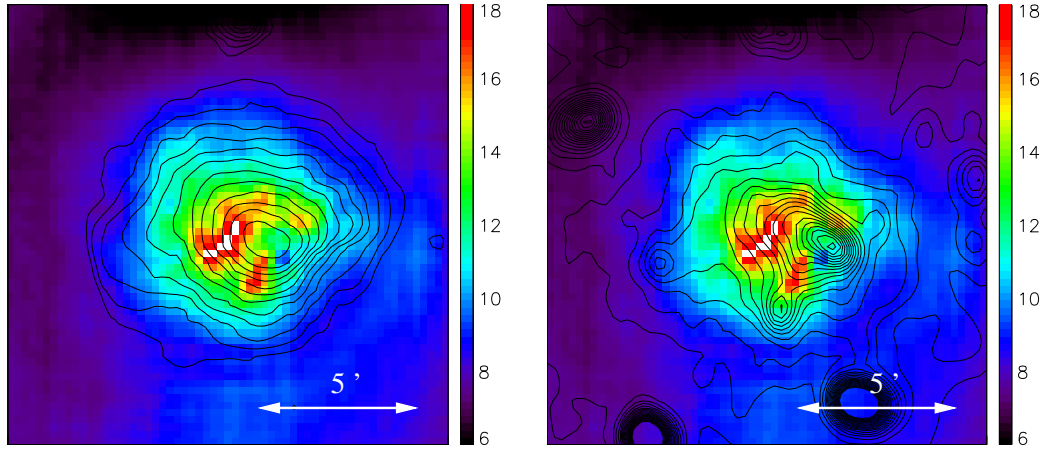


FIG. 2.3 – Carte de température de la région centrale de l’amas de galaxies Abell 2163. (1) Superposition la distribution d’émissivité X. (2) Superposition avec la distribution d’émissivité radio (Feretti, 2001.). Pour un décalage spectral de $z = 0.201$ et avec $H_o = 75 \text{ km.s}^{-1}.\text{Mpc}^{-1}$, $\Omega_m = 0.3$, et $\Omega_\Lambda = 0.7$, 5 minutes d’arc correspondent à 927 kpc.

d’amplitude exceptionnelle ($8 < T < 20 \text{ keV}$). La partie centrale de l’amas est entourée d’un grand arc chaud au Nord-Est ($T \simeq 20 \text{ keV}$), tandis qu’une petite région très froide ($T \simeq 8 \text{ keV}$) est observable au Sud-Ouest, au-dessus de la région de compression des isophotes de densité de gaz (cf. fig. 2.3, carte (1)). Ces résultats sont en accord avec ceux obtenus par Markevitch (2001) avec le satellite Chandra. Disposant d’une meilleure résolution spatiale en émissivité, les auteurs ont mentionné la possible corrélation de la région froide avec un pic d’émissivité X. Ils ont interprété cette structure comme un coeur froid issu d’un sous groupe récemment accréte. Quant à l’arc chaud, il s’agit probablement d’une onde de choc radiale. Notons que sa contribution à l’élévation de la température moyenne de l’amas est très importante. Les auteurs ont par ailleurs mentionné que la morphologie de l’arc chaud était fortement corrélée avec celle du halo radio observé par Feretti (2001), ce que nous confirmons (cf. fig. 2.3, carte (1)). Les halos radios sont en effet supposés être dus au rayonnement synchrotron d’électrons relativistes accélérés par les processus de chocs dans les amas massifs.

Contrairement au cas des amas en fusion bimodale A3921 et A1750, il est difficile d’établir un scénario de fusion pour l’amas A2163 parce que les sous-groupes ont déjà fusionné et sont alors difficilement discernables. Si une onde de choc et un coeur froid sont clairement mis en évidence, la géométrie du problème est difficile à établir à cause des effets de projection et du peu d’informations disponibles concernant la distribution des galaxies dans la région centrale.

Ces résultats EPIC-MOS sont encore préliminaires en particulier à cause de la difficulté à modéliser le fond astrophysique dans cet amas. Pratt et al (2001) ont en effet montré que la composante de fond diffus était très élevée dans cette région du ciel et influait beaucoup sur l’estimation de la température moyenne de l’amas sur les bords de champ. Nous n’avons pas discuté de la forte décroissance en température visible à grand rayon sur la carte (2.3), parce que nous ne sommes pas encore en mesure de la certifier avant d’avoir intégré d’autres observations sur les bords de champ.

Conclusion

Dans le but d'étudier les propriétés du milieu intergalactique dans le cadre de la formation des amas et des groupes de galaxies, nous avons présenté des algorithmes d'imagerie et un algorithme de spectro-imagerie destinés à restaurer la distribution spatiale des paramètres physiques θ du plasma intervenant dans le processus d'émissivité. Ce travail a été appliqué aux spectro-imageurs EPIC équipant le foyer du télescope XMM-Newton.

Pour restaurer la distribution spatiale d'émissivité du plasma, $\theta = L_X$, nous avons comparé divers algorithmes adaptatifs de débruitage et discuté des avantages des algorithmes utilisant la transformée en ondelettes et l'analyse spatio-fréquentielle: décomposition du signal en ses variations, possibilité de construire des algorithmes redondants et itératifs. Nous avons finalement adopté un algorithme spécialement adapté aux processus de Poisson à faible statistique, dans lequel la distribution spatiale du paramètre de Poisson est restaurée en sélectionnant les coefficients d'une transformée en ondelettes de Haar dont l'amplitude dépasse un seuil de signification. Ce seuil est déterminé en fonction de la densité de probabilité théorique des coefficients associés à un processus de Poisson stationnaire (Jammal & Bijaoui, 2000). Le calcul de densité de probabilité a été rendu possible grâce à l'utilisation de l'ondelette de Haar qui permet d'implémenter facilement une dualité entre une statistique de comptage multirésolution dans l'espace direct et une analyse variationnelle dans l'espace temps-fréquence.

Cette expertise nous a conduit à construire un algorithme de spectro-imagerie destiné à restaurer les distributions spatiales d'entropie et d'émissivité du plasma intra-amas. Sur le même principe que dans l'algorithme de Jammal & Bijaoui, on commence par estimer localement les paramètres à partir d'une statistique de comptage multirésolution dans l'espace direct, puis on restaure les variations spatiales de l'estimateur en sélectionnant les coefficients en ondelettes de Haar dont l'amplitude dépasse un seuil de signification vis-à-vis de la densité de probabilité des coefficients associés à un processus stationnaire. Ici, cette densité est supposée gaussienne et sa variance est déterminée en fonction de la variance de l'estimateur. Cet estimateur utilise la méthode du maximum de vraisemblance et se comporte comme un estimateur optimal dans le domaine d'utilisation de l'algorithme. Il prend en compte la réponse instrumentale locale de l'instrument, ainsi que la nature multicomposante du modèle d'émission. L'algorithme a été testé sur des observations simulées proches des conditions réelles, en particulier à l'aide d'une simulation numérique d'amas en cours de coalescence. Ces résultats ont prouvé son aptitude à mettre en évidence de manière satisfaisante les principales structures caractéristiques attendues dans les amas en coalescence, –fronts de choc, coeurs froids, gradient de température centre-bord à grande échelle–, et ce malgré le choix discutable de n'estimer qu'une seule valeur de température du plasma malgré les effets de projection.

Cet algorithme a permis d'étudier les distributions de température et d'entropie de deux amas bimodaux en cours de coalescence (Abell 3921 et Abell 1750). Dans chacun d'eux, nous avons clairement mis en évidence des fronts de choc de compression à l'interface des sous-groupes ainsi qu'une structuration en température et en entropie perturbée dans l'un des sous-groupes. L'analyse de ces cartes à deux dimensions, jointe à celle des profils de densité et de température de chaque sous-groupe à une dimension, a permis d'élaborer un scénario probable de coalescence pour chaque amas (Belsole et al., 2004). Nous avons aussi cartographié la distribution de température de l'amas massif Abell 2163 et mis en évidence de très fortes variations de température ($8 < kT < 20 \text{ keV}$), en accord avec l'hypothèse selon laquelle cet amas est en train de subir ou vient de subir un processus de coalescence violent.

Les perspectives d'évolution de cet algorithme de spectro-imagerie sont multiples. En ce qui concerne l'imagerie, un rééchantillonnage de l'espace adapté aux fortes variations d'émissivité caractéristiques du plasma intra-amas permettrait de mieux explorer l'espace spatio-fréquentiel. La déconvolution de la fonction d'étalement de l'instrument pourrait être envisagée. L'utilisation d'autres ondelettes que l'ondelette de Haar permettrait de réduire quelques artefacts. En ce qui concerne la spectroscopie, l'implémentation d'un estimateur de métallicité du plasma est envisagée. Par ailleurs, l'utilisation de techniques de décomposition du modèle d'émission en composantes principales pourrait permettre de réduire sensiblement le temps de calcul. Cette dernière possibilité a été explorée par Churazov et al. (1996), qui ont montré que l'on pouvait modéliser les spectres d'émission du plasma intra-amas par une combinaison linéaire de deux spectres avec une erreur

relative de 2-4 % dans le régime où l'émissivité du plasma est dominée par le rayonnement de freinage.

A travers les applications étudiées, nous avons montré que les algorithmes multi-échelles sont particulièrement adaptés à l'étude des processus physiques variés intervenant dans le plasma intra-amas. Le caractère automatique de ces algorithmes devrait permettre de généraliser leur application sur des relevés de groupes et d'amas de galaxies de différentes masses et de différents âges et d'étudier statistiquement l'évolution des propriétés globales du plasma au cours du processus de formation des structures. Si l'application d'un algorithme multiéchelle est particulièrement adaptée aux spectro-imageurs EPIC du satellite XMM-Newton dont le champ de vue de 30 minutes d'arc est souvent bien plus large que les détails étudiés, un algorithme similaire devrait aussi être implémenté au foyer du télescope Chandra, qui bénéficie d'une meilleure résolution spatiale.

Quatrième partie

Annexes

Annexe A

Publications.

The detection of diffuse emission in HCG 16 with XMM-Newton

E. Belsole¹, J.-L. Sauvageot¹, T. J. Ponman², and H. Bourdin³

¹ Service d'Astrophysique, CEA Saclay, L'Orme des Merisiers, Bât. 709, 91191 Gif-sur-Yvette Cedex, France

e-mail: ebelsole@cea.fr; jsauvageot@cea.fr

² School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

e-mail: t.j.p@star.sr.bham.ac.uk

³ Observatoire de la Côte d'Azur, BP 4229, 06304 Nice Cedex 4, France

e-mail: bourdin@obs-nice.fr

Received 28 January 2002 / Accepted 10 October 2002

Abstract. We report results obtained from analysis of the XMM-Newton observation of the compact group of galaxies HCG 16. It is a peculiar system composed of 7 spirals, 6 of which are active, and its nature as a bound system has been much debated. The EPIC camera observations give new insights into the X-ray parameters describing the physical status of the group. We detect diffuse X-ray emission with a rather elliptical morphology which extends to at least a radius of $135 h_{50}^{-1}$ kpc from the group centre. The spectrum within this region is well modelled by a thermal plasma with a temperature of 0.49 ± 0.17 keV, and a non-zero metallicity. We measure a bolometric X-ray luminosity of $9.6 \times 10^{40} h_{50}^{-2}$ erg s⁻¹ which may be only a small fraction of the total luminosity because of the limit in spatial detection arising from the high background level. Despite its low temperature and luminosity, HCG 16 obeys the $L_X - T$ relation obtained for brighter galaxy groups even if it lies in a very extreme position. The properties of the diffuse emission confirm the bound nature of HCG 16 even if the gas trapped in the potential well may not yet be virialized. This reopens the debate about the real nature of spiral-dominated galaxy groups, and on their role in a more general cosmological context.

Key words. galaxies: clusters: general – galaxies: clusters: individual: HCG 16 – X-ray: galaxies: clusters – X-rays: general

1. Introduction

Small groups of galaxies are the most common configurations of aggregated systems in the Universe (Tully 1987 and references therein) and they continue to be the subject of intense study. Despite much attention in the last decade, there is continued debate as to whether these systems are genuinely gravitationally bound (Hickson & Rood 1988) or if they are occasional alignments along the line of sight belonging to more extended systems (Mamon 1986), clusters (Walke & Mamon 1989), or cosmological filaments (Hernquist et al. 1995). In particular the interest in compact groups of galaxies, of which the most widely studied catalogue was compiled by Hickson (1982), revealed a series of peculiarities as compared to galaxies in clusters or to isolated galaxies of the same type (see Hickson 1997 and Mulchaey 2000 for reviews).

Galaxy groups have been observed in the X-ray since the advent of the *Einstein* satellite. In fact, the detection of hot gas trapped in the group potential well represents strong proof of the bound nature of these systems.

Several works based on ROSAT/SPSP and ASCA data revealed a variety of morphological and dynamical behaviour in galaxy groups (Pildis et al. 1995; Saracco & Ciliegi 1995;

Ponman et al. 1996 (hereafter PBEB); Helsdon & Ponman 2000b; Mulchaey & Zabludoff 1998).

More than 50% of the 32 ROSAT/SPSP-observed compact galaxy groups (pointed observations) analysed by PBEB exhibit diffuse X-ray emission with temperatures lower than 1.5 keV: the authors found that correcting for the selection effects, the fraction of HCGs with L_X (bolometric) $> 10^{41.1} h_{50}^{-2}$ erg s⁻¹ was 75%. This high fraction ruled out the hypothesis that groups are aligned configurations within cosmological filaments.

Galaxy groups show flatter surface brightness profiles and steeper “luminosity-temperature” ($L_X - T$) and “velocity dispersion-temperature” ($\sigma - T$) relations compared to clusters. This behaviour becomes even more noticeable for very-low temperature ($kT < 1$ keV) systems (Helsdon & Ponman 2000a; Mulchaey 2000; Xue & Wu 2000) where the departures from the self-similar scaling laws traced for galaxy clusters have been explained by models of pre-heating of the gas before it falls into the group (Cavaliere et al. 1997; Ponman et al. 1999; Balogh et al. 1999; Dos Santos & Doré 2002).

These low temperature systems are puzzling for other reasons. Whereas it is well established that a gravitationally bound intergalactic medium is a common property of spiral-poor groups, the results are still controversial for groups composed

Send offprint requests to: E. Belsole, e-mail: ebelsole@cea.fr

of more than 50% of spiral galaxies (Pildis et al. 1995; Saracco & Ciliegi 1995; Mulchaey et al. 1996), which accounts for the majority of groups.

A rare opportunity to investigate these questions is presented by an extreme example of a very nearby ($z = 0.013$) compact group: the Hickson Compact Group 16 (hereafter HCG 16). This group is peculiar because it is composed of 7 galaxies (4 originally detected by Hickson (1982), with 3 more added in a circle of $15'$ by Riberio et al. (1996) on the basis of radial velocity measurements), of which *all* are spiral and 6 of the 7 are active. It is one of the densest groups in the Hickson catalogue, with a mean velocity dispersion of $\sim 100 \text{ km s}^{-1}$ and a median projected galaxy-galaxy separation of $88 h_{50}^{-1} \text{ kpc}$ (Hickson et al. 1992)¹.

Bachall et al. (1984) observed HCG 16 with *Einstein*, determining an X-ray luminosity of $2(\pm 1) 10^{41} h_{50}^{-2} \text{ erg s}^{-1}$, but with the spatial resolution of the instrument the authors were unable to separate the emission of the galaxies from a more diffuse emission.

Using the ROSAT/SPSP, HCG 16 was later studied by Saracco & Ciliegi (1995), who suggested that the X-ray emission was mainly due to the galaxies. However PBEB and Dos Santos & Mamon (1999, hereafter DSM99), analysing the same ROSAT/SPSP data set, found diffuse X-ray emission far away ($\sim 200 \text{ kpc}$) from the galaxies and with a temperature of the order of 0.3 keV , which makes HCG 16 the coolest compact group yet detected.

The detailed spatial analysis of DSM99 ruled out the possibility that the hot gas is virialized. They further suggested that the contribution to the gas enrichment from galactic winds could be an important test to understand the dynamics and evolutionary history of this group.

In this work, we present a detailed analysis of the XMM-Newton observation of the HCG 16, focusing on detection of the hot X-ray gas (we do not discuss the galaxy emission, see Turner et al. 2001a). At the redshift of the group, $1'$ is equivalent to 22.5 kpc .

2. Observations and data reduction

2.1. Observations

The XMM/EPIC (2 EMOS and 1 EPN camera - Turner et al. 2001b; Holland et al. 1996; Strüder et al. 2001) first light observation was a 65 ks exposure of HCG 16. This consisted of two separate exposures (only 1 for EPN) for which each camera was turned on individually (there was no standard mode). Information on the roll angle is not available for these very early data, but we were able to superimpose the two MOS camera image obtained in raw data, and no evident roll angle variation is observable. The observation log is listed in Table 1.

A first data processing has been done with a preliminary version of the SAS (Science Analysis Software) pipeline scripts EMCHAIN (EMOS) and EPCHAIN (EPN). Because of the early stage of the experiment the data files are not in standard configuration, and thus we are unable to process the data with the current (5.3 version) of the SAS software.

¹ throughout this paper we assume $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$; h_0 is the Hubble constant in units of $50 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

Consequently the data are treated using software developed in Saclay for calibration purposes. The output products are consistent with the SAS for downstream compatibility. The processing takes into account the electronic noise cleaning and the difference in exposure time per CCD. This data pre-processing eliminates most of the low energy noise, but from a conservative point of view we limited all analysis to energies above 0.2 keV . A higher noise is detected in some CCDs (CCD2 and 5 for EMOS 2 and CCD 4 for EMOS 1), likely due to the relatively high temperature of some CCDs at this stage of the experiment (they were later cooled). For example, the threshold for telemetry transmission was higher for the CCD2 of MOS2, representing an additional source of noise. Furthermore, this CCD was turned off for one of the EMOS 2 exposures, decreasing the quality of the data from this spatial region.

High background time intervals (due to solar flares) were rejected, for the EMOS cameras, by a careful examination of the 10 to 12 keV light curve. The drop of the EMOS effective area at high energy allows us to be sure that the light curve variation is not due to source variability in this energy band. Events were grouped in 100 s bins and all bins with more than 18 counts were rejected. After this filtering, the useful exposure time is 45 ks.

The EPN observation was made with the very first on-board instrument configuration, which is rather different from that used currently, and which has not been yet calibrated. Furthermore, the available blank-sky background observations were obtained in the standard configuration setting. Because of these problems, we have not attempted to use the PN data for our analysis, restricting ourselves to the 45 ks of useful EMOS data.

Figure 1 shows the EMOS 1 and EMOS 2 summed count rate image of the HCG 16 pointing in the energy band $[0.2-7.0] \text{ keV}$, smoothed with a Gaussian filter of $\sigma = 12''$. The image is not corrected for vignetting. The four main galaxies are marked as in Hickson (1982).

2.2. Background estimate

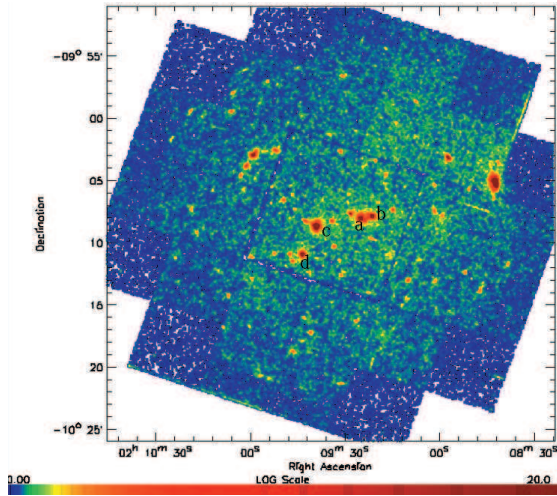
The estimate of the background level is a crucial point since we are interested in a very low surface-brightness, diffuse source. Because of the high background level we can assume that the contribution of any diffuse gas emission at large radii from the group centre is negligible. This allows us to make an estimate of the background in the same field (i.e. a local background). We consider that the XMM-Newton background is composed of two main components:

1. an astrophysical X-ray background which is the combination of a soft ($E < 1.5 \text{ keV}$) component, mainly due to the local bubble, and a hard component due to unresolved cosmological sources (mainly AGN);
2. an induced cosmic ray background which dominates at high energy ($> 5 \text{ keV}$) and induces fluorescence lines (Al, Si, Cu, Au) from the shielding of the camera and the detector itself.

The two components deposit an indistinguishable signal on the detector, but have to be considered as completely different in the data analysis. The first component represents the X-ray

Table 1. The observation log. The net exposure is the sum of exposure 1 and 2 and after flare rejection.

	EMOS 1	EMOS 2	EPN
exposure 1 (s)	55 203	48 302	32 399
comments	50% flare rejection	CCDs 2 and 5 noisy 50% flare rejection	many bright pixels 40% flare rejection
exposure 2 (s)	16 910	17 137	0
comments	—	CCD2 off; CCD5 noisy	—
net exposure (ks)	45 905	43 858	12 525
useful data (ks)	45 905	43 858	0

**Fig. 1.** HCG 16 EMOS 1+EMOS 2 image in the energy band 0.2–7.0 keV. The 4 main galaxies are labelled as in Hickson (1982).

photons that are collected by the telescope and focused on the camera, and so they are affected by vignetting. Conversely, the 2nd component is not affected by vignetting because the particles pass throughout the instrument as a whole. This means that to optimise the analysis of the X-ray background it is better to estimate the particle background before the application of the vignetting correction (see Belsole et al. 2001 for details).

EMOS detections outside the field of view (FoV) of the telescope are only due to particle events. One observation is not sufficient accurately to estimate the particle background, due to the small number of detected events. We thus summed the performance verification (PV) phase EMOS observations to achieve a total exposure time of 2×10^6 s, and considered the out-of-FoV events. These events are detected by all but the central CCD, which has no out-of-FoV region. The spatial and spectral distribution of this particle emission was extended into the FoV, for each CCD, using a Monte Carlo simulation under the following assumptions:

- The particle spectrum outside and inside the FoV are identical:
we use the 2×10^6 s PV observation. The only spectrally

useful region to compare directly the particle spectrum inside and outside the FoV is above 10 keV. In this spectral range, the region inside the FoV is essentially source-free due to the drop in effective area. For each CCD we verified that the spectrum above 10 keV is the same for the outside and inside FoV regions, except for a normalisation factor (of the order of 10%) due to the shielding of the telescope. It is thus an acceptable hypothesis to assume the equivalence between the continuum spectrum inside and outside the FoV in the whole energy band of EPIC.

- It has a uniform spatial and spectral distribution in each CCD:

we looked for variations between CCDs in the same energy band above 10 keV. Examination of the spectrum of each CCD shows that there is some variation corresponding to the fluorescence lines of gold (11.47 keV, present only in CCD2 and 7 for EMOS 1, for example), but that the spectrum of the continuum is very similar from CCD to CCD, only showing a variation of about 5%. We thus estimate the particle spectrum and spatial distribution of the central CCD (which has no out of FoV regions) from the median between CCD 3 and 6, which are closest to it following the camera geometry. Using this approach, we are able to accurately reproduce the high energy (>10 keV) spatial variations due to gold fluorescence lines because at these energies there are no source contributions. We note however, that the Al (1.48 keV) and the Si (1.74 keV) fluorescence lines are not spatially constant. Reproduction of the distribution of these lines is severely complicated by source contributions, and will require careful analysis of CLOSED² data, a statistically useful quantity of which has only recently been released. We take a conservative approach because of the relative weakness of our source, and exclude data from 1.4 to 2.0 keV (fluorescence Al and Si lines) in all subsequent analysis.

- The spectrum of the particle background does not change significantly in time.

While the average induced cosmic ray background level has been shown to be constant within $\pm 10\%$ (see e.g., Arnaud et al. 2001), once periods affected by solar flares are excluded, no significant spectral variation has yet been detected.

² Observations obtained during the calibration phase using an aluminum filter of 1mm thickness, which stops the X-ray photons but not the particles.

In conclusion, the continuum spectrum of the particle background inside the FoV can be well represented using the out of FoV distribution, extended into the FOV using the Monte Carlo method under the assumptions listed above.

EMOS 1 and EMOS 2 particle event lists, with a total simulated exposure time of 10^6 s, were generated in the same format as real EMOS observations and were used for the subsequent image and spectral analysis. As the analysis of the particle distribution was done on the out of FoV events and the flux in this region is lower than in the FoV because of the shielding of the camera, the particle count rate and the source count rate were normalised in the 10 to 12 keV energy band.

3. Analysis and results

3.1. Image and preliminary spectrum

As a first approach in detecting any diffuse X-ray emission, we smoothed the 0.3 to 7 keV EMOS image to find the spatial region where the group diffuse signal is mainly located. To optimise the S/N of the image we performed a preliminary spectral analysis by extracting a spectrum in a circle of $6'$ in radius around the optical centre of the group (i.e. where the smoothed image shows a significantly denser emission), excluding the signal from the galaxies. We subtracted from it the spectrum of the $10'-12'$ ring in order to have an estimate of the local background. This rough analysis allows us to calculate an upper limit to the flux and to determine that the peak of the X-ray diffuse emission is around 1 keV (as previously shown in ROSAT data). Also, no emission is detected above 2 keV where the spectrum is dominated by the emission from individual galaxies (see Turner et al. 2001a). Because this work is focused on the group diffuse X-ray emission, unless otherwise stated we limit the following analysis to the soft energy band [0.2–1.4] keV, the upper energy limit being determined by the contamination of the Al and Si fluorescence lines (see Sect. 2.2).

3.2. Wavelet analysis

The wavelet de-noising is a multi-scale image reconstruction tool. The noise contribution to the signal is removed by selecting its corresponding coefficients in the wavelet space. A new iterative wavelet de-noising algorithm (based on the *à trous* algorithm) especially suited for images with few counts is used here. Within this algorithm, thresholds are computed analytically for a Poisson noise distribution (see Bourdin et al. 2001). This algorithm has been written to be used for restoring the surface brightness of the X-ray emitting intra-cluster gas which is supposed to have no steep discontinuities. In the case of a group like HCG 16, the individual galaxy emission introduces a high level of discontinuity with a dynamic of three orders of magnitude. A crude application of our algorithm to such data would be good at restoring the emission of the extended gas but the restored image would be contaminated by artifacts around these highly emitting sources.

In order to exclude this contribution, a count threshold was first defined on the image, and the bright sources identified.

Then the flux inside the source emitting regions, as well as in the CCD gaps, was interpolated so as to create a new image without steep discontinuities. The image was de-noised with the restoration algorithm and the sources were then re-added.

We have defined a counts (Poisson's statistic, i.e. no vignetting correction is done) image of $4'1$ pixel size where the contributions of the two EMOS cameras are added. The energy band 0.2–1.4 keV has been chosen so as to be sure that the detection excludes the Al and Si fluorescence lines. An image in the same energy band is also obtained for each camera, from the particle event lists. The EMOS particle images are then coadded.

The HCG16 counts image is de-noised with the previously described wavelet de-noising procedure. The threshold probability adopted for the HCG 16 image is of 10^{-6} , meaning that there is a probability of 10^{-6} that the detected structure comes from noise, or in other terms, the structures detected in the image are significant at 4.5σ . The adopted threshold is quite severe, but due to the low S/N we prefer to adopt such a conservative approach. We detect significant extended emission only at scale 6, corresponding to $4/4$ at the pixel size of the image. We do not detect any extended emission at higher frequency scales, indicating that if more clumpy structures exist, these cannot be significantly detected with these data. We are unable to distinguish between the real galaxy extension, the effect of the Point Spread Function (PSF) wings and the actual diffuse emission. However, it is unlikely that at the scale of detection the signal comes from the galaxies only (see Sect. 3.4.2).

The particle background image is smoothed with a Gaussian kernel of $\sigma = 1/4$, comparable to the scale of the large structures in the wavelet image. This operation allows us to take into account boundary effects which would be ignored if we just subtracted a constant value per CCD. The smoothed particle image was then normalised by the exposure time and subtracted from the denoised counts image. We then corrected the image for vignetting and the net result is shown in Fig. 2. To aid in the interpretation of the observed features, we superimpose on this figure the CCD gaps.

The wavelet detection of diffuse gas between galaxies exhibits an elliptical shape which extends up to a radius of $\sim 6'$ (~ 135 kpc) from the optical centre of the group to the NW-SE direction and to $4/5$ in the other direction.

However, there are some noticeable effects in the image (we show in Fig. 3 the position of CCDs in the HCG16 sky coordinates):

- The two noisy CCDs of EMOS 2 (CCD2 and CCD5) have not been excluded during the wavelet analysis in order to have a more continuous image. On the one hand, as our electronic noise cleaning is not perfect, the final result is an artificially high signal in the corresponding regions (see Fig. 3). This explains the high flux to the north-west (CCD5; note how this higher emission follows the shape of the CCD) and to the south-east (CCD2). On the other hand, for one exposure, CCD2 was turned off, and thus the total (EMOS 1+EMOS 2) image has a lower S/N ratio in the south-east zone corresponding to the EMOS 2 CCD2. The net effect is that, in this region, the structure detection will

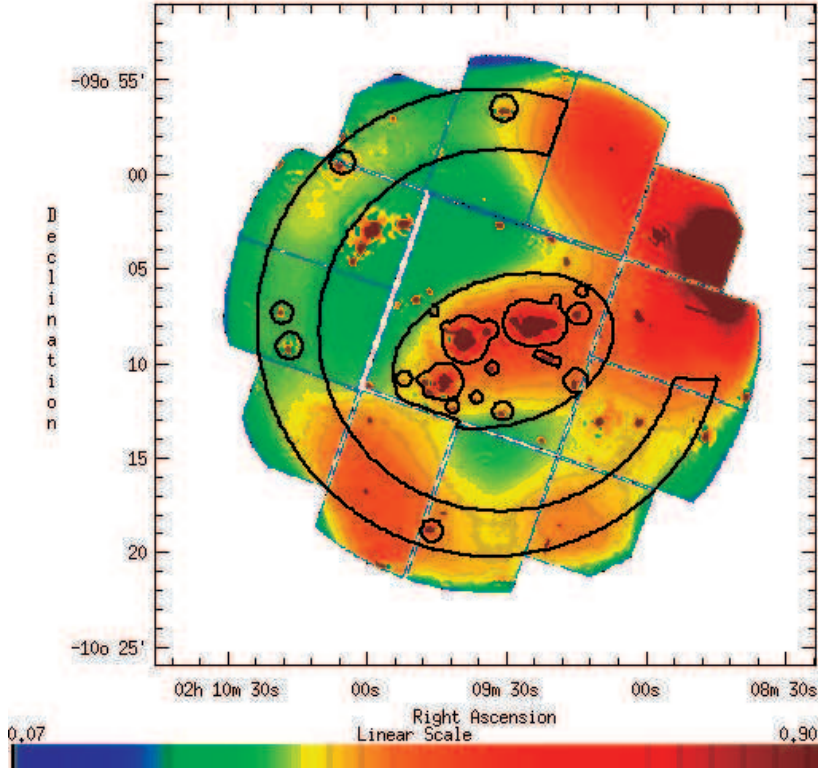


Fig. 2. EMOS 1+EMOS 2 wavelet image in the energy band 0.2–1.4 keV. The ellipse at the centre of the figure denotes the region used for spectrum extraction. The ring from which the background spectrum was extracted is also shown to indicate the distance from the elliptical region. A slightly different ring is used for EMOS 1 and EMOS 2 in order to exclude the noisy CCDs.

be less efficient (our wavelet algorithm does not allow for the moment to give a different weight to different spatial regions), and thus the smoothing kernel will appear larger (in this region) in the reconstructed wavelet image. Discarding the noisy CCDs would noticeably increase this effect.

- The high signal to the west-north-west is further increased by a point source highly degraded by the off-axis PSF;
- The extension onto the south-east (in the direction of galaxy d) is also probably due to a radio source (DSM99). This source is detected on CCD2 of EMOS 2, but because of the high CCD temperature in the early stage of the calibration, there is a great deal of electronic noise which effectively prevents further investigation of the source.

3.3. The radial profile

The gas density distribution, obtained from the surface brightness profile, is one of the parameters (together with the temperature and total flux) used to measure the gravitational potential well in the assumption of spherical symmetry.

We define a mask excluding all sources detected in the wavelet image as well as detector artifacts such as bad

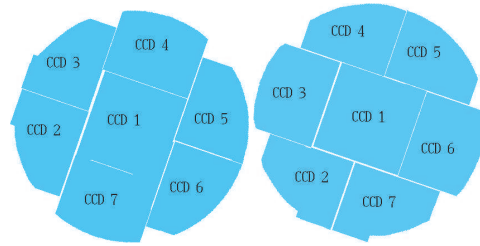


Fig. 3. Position of the CCDs for each EMOS camera (EMOS 1 to the left, EMOS 2 to the right) with the orientation of the HCG1 sky coordinates; north is up the east is to the left.

lines/columns and hot CCDs not taken into account by the pre-processing. Because of the uncertainties in the spatial distribution at the energies corresponding to the Al and Si fluorescence lines in the particle image (see Sect. 2), we again limit our analysis to the 0.2–1.4 keV energy band. We are confident that this choice does not change our result because the preliminary

spectral analysis (Sect. 3.1) showed that the X-ray diffuse emission has its maximum around 1 keV.

To correct for vignetting, we used the photon weighting method described in Arnaud et al. (2001): each photon is weighted by the ratio of the effective area at its position on the detector to the central effective area at the energy of the photon. To take into account the “false” vignetting correction of the particle background component, we also corrected the particle image for vignetting, in the same energy band, and with the same algorithm. In the energy band considered here, the contribution of the particle background is, in principle, negligible. However, because of the telescope vignetting, in the outer regions more than half of the X-ray photons are lost, while the particle flux stays constant across the whole FoV. Thus the particle background contribution must be subtracted because, as shown in Fig. 4, at large radius ($>7'$) this component contributes $\sim 50\%$ of the total sky background.

The particle image was normalised, as described in Sect. 2, to the 10 to 12 keV energy band count rate. This was done separately for EMOS 1 and EMOS 2 and only after normalisation were the two images added together.

The surface brightness profile was obtained for the normalised (EMOS 1+EMOS 2) particle image and for the HCG 16 image separately (Fig. 4a), but using the same mask to exclude point sources. Afterwards, the particle radial profile was subtracted from that of HCG 16 (Fig. 4b). To summarise, if $I = E \times S + P$ where I , S , P and E indicate the observed image, the source image, the particle image and the exposure map (including vignetting), respectively, then the whole radial profile analysis can be written as $S = I/E - P/E$.

An estimate of the mass of the diffuse intergalactic medium is given by fitting the radial profile with a King profile (β model). However the HCG 16 gas shape (Fig. 2) indicates that the β -model method assumption of spherical symmetry is not necessarily true. The centre of the fit was obtained, after several iterations, in order to maximise the central X-ray surface brightness of the King profile. The best solution was $\alpha = 02^h09^m32.0^s$; $\delta = -10^\circ09'00''$ (J2000), which corresponds within $0.3''$ to the optical centre (given by SIMBAD, FK5 coordinates). The best β -model fit of the surface brightness profile gives a β parameter of 0.37, a core radius of $2.7'$ and the central intensity of 6.5×10^{-4} cts s^{-1} arcmin $^{-2}$ (Table 2). To check the consistency of this result and to have an estimate of the statistical significance of the error bars (given at 1σ confidence level), we determine the background level independently, i.e. computing the count rate in the $[0.2-1.4]$ keV energy band and in the $10'$ to $12'$ ring (we consider it safer to limit the large radius analysis to $12'$ since outside this radius some systematic error may come from the vignetting function, which is poorly known beyond this radius). This gives a value of 6.0×10^{-4} cts s^{-1} arcmin $^{-2}$ which is consistent with the value found when we fit the profile with the background parameter left free. However, due to the very low S/N , the fit is very sensitive to the background level. This may have an important effect on the fitted parameters, and also in the mass determination. The poor constraints on the fitted parameters are thus not surprising and we have to be careful with regard to their use in the determination of physical properties.

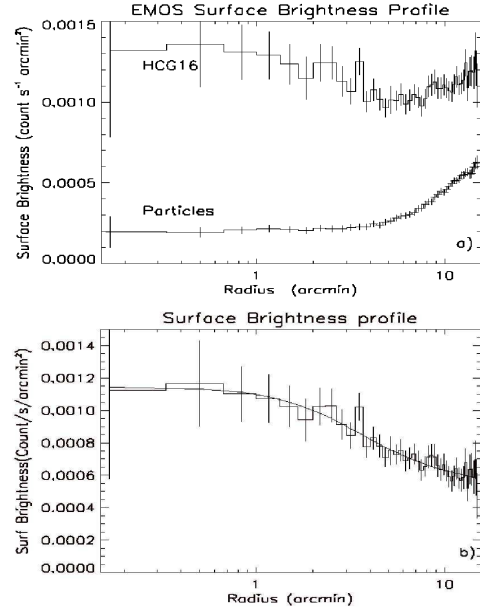


Fig. 4. **a)** The surface brightness profile for the HCG 16 (EMOS 1+EMOS 2) and for the simulated particle image in the 0.2–1.4 keV energy band. Both profiles have been calculated from images corrected for vignetting. The two radial profiles are computed using the same mask and in the same energy range. **b)** The EMOS 1+EMOS 2 surface brightness profile after subtraction of the particle background. The parameters of the β model fit are listed in Table 2.

3.4. The spectra

3.4.1. Method

While a standard technique for extracting spectra is to consider a circle around the source centre, we use a different method to determine the spatial extraction region due to the shape of the detected signal (Fig. 2), the low surface brightness (Fig. 4b) and the low S/N of the observation. Assuming that the wavelet analysis is the best representation for the detection of the diffuse emission, we define a regular elliptical region, following the shape of the emission in the wavelet image (Fig. 2). Within this elliptical mask, all bright sources have been excluded by defining a threshold (1 cts/s) in the wavelet image, which results in a typical radius of $1.1'$ for the exclusion of the galaxies. We will refer to this elliptical mask region as region E .

Given that the intragroup gas is detected mainly in the central CCD (Fig. 2), it is possible to estimate a local background in an external region corresponding roughly to a $10'$ to $12'$ ring around the centre of the FoV (Fig. 2). We will refer to this local background as region R .

To avoid the introduction of systematic errors, we are, in this spectral analysis, more conservative than in the imaging analysis and in both regions E and R we exclude: 1) the south-east corner because of contamination due to CCD2 of EMOS 2, which has a much higher noise (see Sect. 2); 2) the region corresponding to CCD5 of EMOS 2, the other CCD displaying

Table 2. The β model fit results. Errors are quoted at 1σ significance.

Parameter	fit value
β	0.37 ± 0.3
R_c (arcmin)	2.7 ± 2.5
I_0 (cts s ⁻¹ arcmin ⁻²)	$(6.5 \pm 0.9) \times 10^{-4}$
bkg (cts s ⁻¹ arcmin ⁻²)	$(5.0 \pm 3.0) \times 10^{-4}$

higher noise and which shows excess emission at low energies; 3) in both cameras, the region to the WNW because of a highly PSF degraded point source (note that this region corresponds essentially to CCD4 in EMOS 1).

The false vignetting correction of the particle component of the local background will over-subtract the true background contribution. To avoid this error, the events are not initially corrected for the vignetting effect. The method used follows these steps:

1. We extracted 2 spectra from each events list (HCG 16 and particle): the first in region E , the second in region R . We now have spectra E_s, E_p, R_s, R_p where the s spectra are extracted from the HCG 16 event list and the p spectra come from the particle event list.
2. The particle spectra were subtracted from the HCG 16 spectra in each region after a normalisation (n) in the 10 to 12 keV band:

$$E_x = E_s - nE_p$$

$$R_x = R_s - nR_p.$$

This gives the particle subtracted spectra E_x and R_x .

3. *Vignetting correction:* in order to compute the vignetting correction factor (VCF), we work under the assumption that the spatial distribution over the whole integrated region is independent of the energy (i.e. the spectrum is similar throughout the region). We thus use all photons, irrespective of their energy, to describe the spatial distribution of events. Then, for each photon energy, we compute the vignetting factor by averaging the weight (see Sect. 2 for the weight definition) at each photon position. The advantage of this method is that in using the total number of photons to calculate the weight, we can improve considerably the accuracy of the calculated vignetting factor. Spectra E_x and R_x were thus multiplied by their respective vignetting correction factors (VCFs):

$$E_{x,\text{vcf}} = E_x \times \text{vcf}_E$$

$$R_{x,\text{vcf}} = R_x \times \text{vcf}_R.$$

4. The spectrum $R_{x,\text{vcf}}$ was then normalised to the surface area of region E , and subtracted:

$$E_{\text{net}} = E_{x,\text{vcf}} - mR_{x,\text{vcf}}$$

where m is the normalization. This gives the full background-subtracted, vignetting-corrected net spectrum, E_{net} .

3.4.2. PSF considerations

An important point of concern is the contamination of the excised sources due to the wings of the EMOS PSF. The one dimensional EMOS PSF is well described by a King function out to $5'$ radius, and in the energy range of interest here, is essentially energy-independent (Ghizzardi 2001). The asymptotic slope of the PSF model is $\alpha = 1.46$ and $\alpha = 1.41$ for EMOS 1 and EMOS 2 respectively, at 0.8 keV in the on-axis position. We adopt this description of the PSF to correct the spectrum of the diffuse gas for the PSF of the excised sources (galaxies and point sources) in the ellipse and background (ring) regions.

In determining the spectrum in the elliptical region, we have excised the sources using a typical radius of $1.1'$ for the galaxies and $24''$ for the point sources. The sources in the background ring were excised in a cutoff radius of typically $40''$.

For each source we calculated the encircled energy fraction (EEF) at 0.8 keV (corresponding to the peak of the diffuse emission spectrum) in the circle used to excise it, taking into account its off-axis position. The fraction of this energy falling outside the excised circle is $F = (1 - \text{EEF})/\text{EEF}$. Once the fraction for each source was obtained, we computed a statistical mean fraction $\langle F \rangle$ for the two regions of interest, the ellipse and the background. The mean fraction $\langle F \rangle$ is 0.27 for EMOS 1 point sources in the ellipse, for example. This means that 0.27 of the flux of the excised point sources contaminates the diffuse X-ray gas flux of EMOS 1.

The PSF correction for the galaxies is more subtle because, in principle, we have to take into account the extension of a galaxy and convolve the emission distribution with the PSF. We thus use the EMOS PSF model to investigate the systematic errors introduced by treating a galaxy as a point source and as modelled by a β model. We applied this test to HCG 16c, the brightest and most extended galaxy, and thus the most extreme case. We obtained a surface brightness (SB) profile of this galaxy from the EMOS 1 image in the energy range 0.2–1.4 keV and we fitted it with the model of a point source convolved with the EMOS PSF. We fit this profile up to $2'$ after which the profile becomes dominated by the diffuse emission. The fit is not good, but if we extrapolate the model up to $5'$ we obtain a $F = 0.063$ using a cutoff radius of $1.2'$ for this galaxy.

We then obtained a SB profile of galaxy HCG 16c from the Chandra image in the same energy range as above. The Chandra PSF can be considered as perfect with respect to the EMOS PSF; we thus assume that this profile represents the intrinsic profile of the galaxy. We fitted the Chandra SB profile with a β model finding as a best fit $\beta = 0.75^{+0.08}_{-0.06}$, $R_c = 5''.28^{+1''.32}_{-1''.02}$ and a reduced χ^2 of 25.37/25 ≈ 1 .

To estimate the EMOS spread function we fitted the EMOS profile (up to $2'$) with a β model convolved with the EMOS PSF and with the β and the R_c fixed to their respective Chandra best fit values. The central X-ray surface brightness and background were free parameters. In this case we obtain $F = 0.07$. If we leave the β parameter free, we obtain $\beta = 0.81 \pm 0.04$ for a reduced χ^2 of 1.18. This value is in good agreement with the value obtained from the Chandra profile. In this case, the obtained F is 0.068.

Table 3. Spectral fit results before and after PSF correction. *Top:* Col. (1): absorbing hydrogen column density, we adopt the Galactic value from Stark et al. (1992). Column (2): temperature. Column (3): metallicity. Column (4): χ^2 of the spectral fit. Column (5): redshift. *Bottom:* Col. (1): X-ray luminosity computed in the PSPC energy band. Column (2): bolometric X-ray luminosity, uncorrect for absorption. Column (3): bolometric X-ray luminosity, uncorrect for absorption and corrected * for the omitted emission from the galaxy regions (follow the discussion in Sect. 4.1). Column (4): bolometric X-ray luminosity corrected for absorption. Column (5): bolometric X-ray luminosity corrected for absorption and corrected as in Col. (7)*. Errors are at the 90% confidence level. The error on the luminosity has been obtained by fixing the metallicity parameter range in the fit to be between 0 and 2 solar.

no PSF corrected				
N_{H} (10^{22} cm^{-2})	kT (keV)	Z/Z_{\odot}	$\chi^2/\text{d.o.f.}$	z
0.02 (fix)	$0.53^{+0.11}_{-0.13}$	$0.09^{+0.12}_{-0.04}$	13.84/20	0.0132
L_{X} ($10^{41} \text{ erg s}^{-1}$)				
0.2–2.3 keV	bol. uncorr.	bol. uncorr. (*)	bol. corr.	bol. corr. (*)
$0.51^{+0.05}_{-0.07}$	$0.53^{+0.07}_{-0.07}$	0.69	$0.92^{+0.13}_{-0.11}$	1.2
PSF corrected				
N_{H} (10^{22} cm^{-2})	kT (keV)	Z/Z_{\odot}	$\chi^2/\text{d.o.f.}$	z
0.02 (fix)	$0.49^{+0.19}_{-0.16}$	$0.07^{+0.31}_{-0.05}$	6.35/11	0.0132
L_{X} ($10^{41} \text{ erg s}^{-1}$)				
0.2–2.3 keV	bol. uncorr.	bol. uncorr. (*)	bol. corr.	bol. corr. (*)
$0.39^{+0.08}_{-0.08}$	$0.40^{+0.08}_{-0.08}$	0.52	$0.72^{+0.12}_{-0.12}$	0.96

To summarise, if we consider this galaxy as a point source, we make a maximum systematic error of 10%. This induces an error on the X-ray luminosity obtained from the spectrum of the diffuse gas of $\sim 1.5\%$, which is consistent with the statistical errors. In view of the results outlined above it appears to be justified to treat galaxies as if they were point sources. We computed a mean $\langle F \rangle$ value for all galaxies (per camera) obtaining $\langle F \rangle = 0.073$ for EMOS 1 and $\langle F \rangle = 0.089$ for EMOS 2. The mean cutoff radius was $68''$ for both cameras. In the calculation of the EEf the off-axis position was taken into account for each galaxy; we note however that this does not introduce a strong effect.

We extracted the global spectrum (the best representation of the total flux) of all point sources in the background (PR; ring) region and in the ellipse (PE) region. The global spectrum of the 4 galaxies was also obtained (G). We then calculated a background spectrum B_c which takes into account not only the PSF contamination due to its own point sources, but also that due to the point and extended sources in the elliptical region. In other words:

$$B_c = B - (PR \times \langle F \rangle_{PR}) + (PE \times \langle F \rangle_{PE} + G \times \langle F \rangle_G). \quad (1)$$

3.4.3. Spectral fit results

The EMOS 1 and EMOS 2 spectra were binned to 2σ statistical significance after background subtraction, and fitted with an absorbed MEKAL model (XSPEC v11.0), the column density being fixed to the galactic value ($2.0 \times 10^{20} \text{ cm}^{-2}$), leaving the temperature, abundance and normalisation (emission measure) as free parameters.

The EMOS 1 and EMOS 2 results are in very good agreement, except in the 0.2–0.3 keV energy band, where the EMOS 2 shows excess emission. We attribute this difference

Table 4. Relative contribution of each galaxy and total point sources to the 22% flux reduction due to the PSF contamination. The contribution of each galaxy has been obtained by integrating the total counts in the cutoff radius (in the energy range 0.2–1.4 keV for EMOS 1 and 0.3–1.4 keV for EMOS 2) used to excise it and by multiplying this value for the mean PSF correction factor $\langle F \rangle_G$. HCG 16a and HCG 16b are considered together. The total point source contribution has been estimated using $\langle F \rangle_{PE}$ as the PSF correction factor.

region	EMOS 1 %	EMOS 2 %
HCG 16a + HCG 16b	23.6	22.7
HCG 16c	23.8	25.5
HCG 16d	7	8.8
point sources (total)	45.5	43

to an instrumental effect not well taken into account in the data processing (probably due to the contamination of the central CCD by the high temperature CCDs nearby, particularly affecting the low energy band) and decide consequently to ignore, when fitting, the 0.2–0.3 keV band in EMOS 2. We used the response matrix version 20.6. In Fig. 5 the EMOS 1 and EMOS 2 spectra and the folded model are shown. In the energy range 0.2–1.4 keV the EMOS 1 camera provides 1635 counts, corresponding to 313 net counts after background subtraction, while the EMOS 2 measured counts are 1269 in the energy range 0.3–1.4 keV (327 net counts). The spectral fit results are listed in Table 3, which also lists the fit results before the correction for the PSF. We can notice that the correction for the wings of the PSF, obtained by using B_c instead of B for the background spectrum (see Eq. (1)), has the effect of slightly reducing the temperature, but above all of reducing the flux by $\sim 20\%$. The relative contribution of each galaxy (treated as a point source) and total point source contribution to this 20% flux reduction is shown in detail in Table 4.

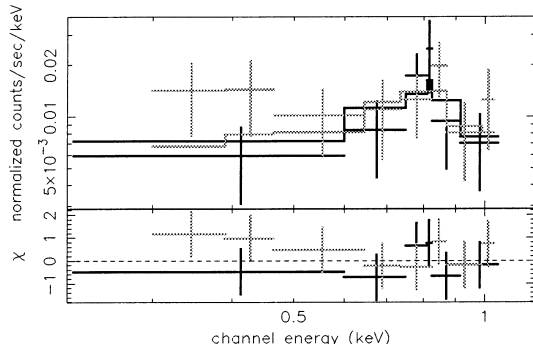


Fig. 5. EMOS 1 (black) and EMOS 2 (gray) background subtracted PSF corrected spectrum, folded with the combined fitted model. The background subtracted spectra are binned to the 2σ level. The channels corresponding to 0.2 to 0.3 keV were ignored for EMOS 2 because of low energy noise not taken into account in the data processing (cf. CCD temperature).

4. Discussion

4.1. Comparison with previous works

The image analysis presented in Sect. 3 and the surface brightness profile confirms that HCG 16 is a bound system as previously shown from ROSAT studies. The EMOS sensitivity is 1.8 times that of the PSPC instrument on-board ROSAT at low energies, and the useful XMM-Newton observation time is more than triple that of the ROSAT/PSPC observation. This allows a better source detection (while point sources are in the noise in the PSPC observation) and exclusion of the galaxy contribution and the establishment of stronger constraints on the temperature. On the other hand, the high background level of XMM (worsened by the electronic noise at the early stage of the instrument setting) reduces the S/N and the detection radius for extended sources. Relative to ROSAT/PSPC, the EMOS PSF wings are larger, leading to a smoother apparent emission than was found by DSM99. In our analysis we have taken into account all these major sources of contamination, finding that the results are stable in the limit of this first light observation. The diffuse emission extends to at least $6'$ from the group optical centre (corresponding to ~ 135 kpc at the distance of HCG 16). Our results show a smooth gas distribution, rather in contrast to the very irregular morphology noted by DSM99. These authors emphasise the nature of region C4, corresponding to an excess emission starting from galaxy HCG 16b toward the south. A quick comparison between the ROSAT/PSPC and the EMOS image shows that the emission to the south-western extreme of our ellipse corresponds to a point source which was marginally detected by ROSAT and which is completely resolved by EMOS. If we smooth the EMOS image in a similar way as done by DSM99, we do not find a spatially continuous signal between the point source (widened by the large Gaussian kernel) and the diffuse emission. We thus estimate that at least 1/3 of the C4 emission is due to a point source unrelated to the group. While the statistics are poor, the spectrum

of this source is minimised by a power law of $\alpha = 2.25$ and a column density of $6.7 \times 10^{19} \text{ cm}^{-2}$. The flux of the source is $1.7 \times 10^{-14} \text{ ergs cm}^{-2} \text{ s}^{-1}$ and if the source was at the HCG 16 redshift its total luminosity would be $1.3 \times 10^{40} \text{ erg s}^{-1}$. The remaining 2/3 of the flux are effectively inside the ellipse in Fig. 2, and there is no clear sign of separation from the neighbouring regions of the diffuse emission. An additional source lies in the region corresponding to C4 in DSM99 (closer to galaxy b). It has been excised in our analysis, but we checked the contribution of this source to the C4 emission. The best fit is a power law of index 2.3, with a flux of $5.3 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$.

The elliptical shape and the spatial extension of the entire gas distribution make it unlikely that this gas can come only from ejection by the galaxies: even if active galaxies (as is the case here) eject large amounts of gas via galactic winds (Fabbiano et al. 1988), this emission does not extend out to several galactic radii.

While HCG 16 still remains one of the coolest galaxy groups with detected diffuse X-ray emission, the EPIC/MOS spectra analysis of HCG 16 gives a temperature of 0.49 keV, higher than that obtained with ROSAT (PBEB, DSM99), but in agreement within statistical errors.

The best fit metallicity is 0.07 solar, which is low with respect to galaxy clusters and some galaxy groups. However, even though several authors have reported lower abundances in groups than in clusters (e.g. Davis et al. 1999), the reality of this result is not well established, and we have to be careful with abundance determinations of such a low temperature systems. The abundance derived from hot plasma model fits for systems with temperatures below ~ 1 keV is primarily driven by the iron L-shell complex, and the combination of the different lines is very sensitive to the temperature. This means that if temperature gradients are present, the abundance is, in general, an underestimate (Buote & Fabian 1998; Finoguenov & Ponman 1999). Works based on ROSAT/PSPC (Mulchaey & Zabludoff 1998) and ASCA (Finoguenov & Ponman 1999) data have already pointed out the multi-temperature structure of several galaxy groups. We are unable to determine a temperature profile for HCG 16 with these data, but given the presence of several starburst galaxies interacting in the central region it seems unlikely that the gas will be isothermal. On the other hand, our best fit metallicity determination is poorly constrained. Moreover the metallicity depends on the model used and it has also been shown, for other groups, that an increase of the metallicity value is obtained if a 2 temperature model is used (Buote 2000). In any case, we find a non-zero metallicity.

The low value of the bolometric X-ray luminosity is not surprising. Galaxy groups, especially at low temperature, have flat surface brightness distributions (Ponman et al. 1999; Helsdon & Ponman 2000a). In this work, the X-ray luminosity is calculated from the emission measure of the MEKAL model used to fit the spectrum. The latter is extracted within the radius where significant X-ray emission is detected, thus the luminosity calculated here may be only a small fraction of the total luminosity.

For comparison, we consider the whole region C2+C3+C4 in DSM99, taking the metallicity as 0.1 solar. Our result ($0.52^{+0.1}_{-0.1} \times 10^{41} \text{ erg s}^{-1}$) and theirs ($0.69 \times 10^{41} \text{ erg s}^{-1}$) do not

strongly disagree if we consider that $\sim 1/3$ of the X-ray emission from C4 comes from a point source, and that the spatial regions and the energy band where the spectrum is extracted are not exactly the same.

In comparison to the PBEB result, $L_X = (4.7 \pm 1.1) \times 10^{41} \text{ erg s}^{-1}$, the bolometric X-ray luminosity found in the present study is more than a factor 5 lower (4.7/0.72). There are several reasons for this. According to DSM99, a factor of 2 of the luminosity found by PBEB is attributable to foreground/background sources, while the good EPIC sensitivity allow us to excise some sources missed in the PSPC analysis by these authors. Our limited ability to detect diffuse emission at large radii due to the high EPIC background is another reason for the difference in the surface used to extract the spectrum, although DSM99 argue that, apart from region C4 there is no emission far from the galaxies. Moreover, in our work we use a different correction for absorption relative to PBEB because of our higher temperature.

Furthermore, we have to consider that in addition to the fact that PBEB used a larger region to extract the spectrum (200 kpc), the flux from the spatial region occupied by the galaxies has been omitted in our analysis, whilst PBEB included an estimate of the diffuse group flux underlying the galaxy components. Hence the total luminosity within the elliptical region considered represents a lower limit to the total diffuse flux of the system. Assuming that the area omitted behind the galaxies has the same mean brightness as the rest of the region, we calculate the ratio between the mask area and the whole elliptical area (Fig. 2). This factor gives a rough correction of 30% to be applied to the L_X value leading to an estimate of $9.6 \times 10^{40} \text{ erg s}^{-1}$ within the full elliptical region. To investigate further the comparison with ROSAT results, we analysed the PSPC data within a $5'$ radius of the EPIC detection. Computing the count rate in this circle (which is comparable to our mask region) and replacing the galaxy emission with the interpolated flux (as done in PBEB), we obtain a bolometric luminosity of $7 \times 10^{40} \text{ erg s}^{-1}$, thus the results from EPIC and ROSAT, within similar central areas, are in an acceptable agreement.

4.2. Mass estimates and scaling laws

An estimate of the total mass of the gas can be obtained by combining the β -model parameters and the spectral results. This implies that we assume, despite our reservations (Sects. 3.4 and 3.3) that the gas is isothermal and has a spherical symmetry. Under these hypotheses, the mass of the gas amounts to $8.8 \times 10^{10} M_\odot$ at the maximum radius of detection (135 kpc). If the mass profile is extrapolated out to the virial radius ($\approx 900 \text{ kpc}$)³ the total gas mass is $3.7 \times 10^{12} M_\odot$. We stress that we have to be careful about this result because we have made several assumptions which may not be strictly valid. Simulations indicate that the β value obtained fitting the surface brightness profile is very sensitive to the range of radii used in the fit (Navarro et al. 1995; Vikhlinin et al. 1999) and

³ the virial radius is obtained as: $r_{\text{vir}} = 4.1 (T/10 \text{ keV})^{0.5} (1+z)^{-1.5} h_{50}^{-1}$ (Evrard et al. 1996).

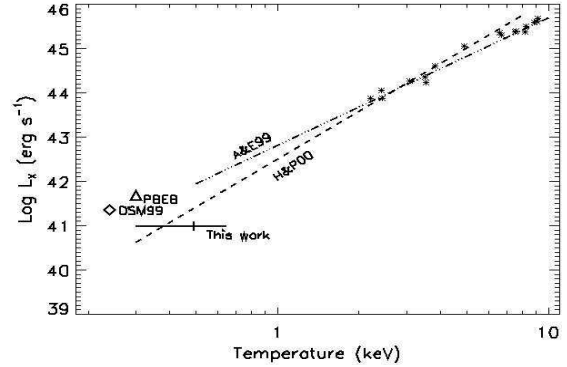


Fig. 6. $L_X - T$ relation for cluster of galaxies (from Arnaud & Evrard 1999 - H&E99) compared to the one for group found by Helsdon & Ponman (2000) (H&P00). Our result is marked as well as the DSM99 (diamond) and PBEB (triangle) results.

in the case of very low surface brightness the background determination is a strong source of uncertainty.

Self-similar scaling laws valid for clusters of galaxies break down at the low mass limit (PBEB; Helsdon & Ponman 2000a; Xue & Wu 2000). In the isentropic limit, semi-analytical models (Cavaliere et al. 1997; Balogh et al. 1999) predict a change in the slope of the $L_X - T$ relation at a temperature around 1 keV, which is in agreement with the observational results from galaxy groups (PBEB, Helsdon & Ponman 2000a,b; Mulchaey 2000). This departure from self-similarity of low temperature systems can be explained by preheating models, in which the gas is heated by some non-gravitational effect (such as galaxy winds during the galaxy formation) before it falls into the potential well of the group (Helsdon & Ponman 2000a; Babul et al. 2002; Tozzi et al. 2000; Dos Santos & Doré 2002). In this context, HCG 16 becomes even more interesting: its low surface brightness, low temperature and low velocity dispersion locate this group in an extreme position in all scaling laws determined for clusters and groups. Figure 6 shows the $L_X - T$ relation from Arnaud & Evrard (1999) for clusters, together with our result, where we take the bolometric luminosity corrected for absorption and for the omitted region corresponding to the galaxies. The dot-dot-dashed line is the best fit for clusters, whereas the dashed curve is the best fit obtained by Helsdon & Ponman (2000b) using the same sample as PBEB (22 groups observed with ROSAT/PSPC) but considering a different statistical weight when fitting the $L_X - T$ relation. HCG 16 is a factor 10 lower with respect to the cluster relation. However, despite its extreme nature, it is in reasonable agreement with the correlation found for galaxy groups, especially if one considers that integration to a larger radius could increase L_X by a factor of a few. One should also bear in mind that the $L_X - T$ relation for groups shows significant scatter (Helsdon & Ponman 2000b), interpreted as arising from the different formation history of individual groups.

Another manifestation of the breaking of self-similarity seen in groups is the progressive flattening in X-ray surface

brightness profiles in cooler systems (Ponman et al. 1999). For a system with a mean temperature as low as that of HCG 16, this trend implies a surface brightness index β less than 0.4 (Helsdon & Ponman 2000a). Our fitted value, $\beta = 0.37 \pm 0.3$, (Table 2) is therefore well-matched to the trend seen in other X-ray bright groups.

Combining our fitted gas density profile with the measured gas temperature (which we take to be isothermal), it is possible to derive an entropy profile for the intergalactic gas. The result is a rising profile, with a value, at $r < 6'$ of $\sim 100 \text{ keV cm}^2$. Although the profile itself should not be taken too seriously, due to uncertainties in the radial trends arising from the limited spatial extent over which we can reliably trace the emission, the entropy derived in the inner regions should be a secure estimate. This is very similar to the “entropy floor” value found in the inner regions of X-ray bright groups by Ponman et al. (1999) and Lloyd-Davies et al. (2000).

4.3. The nature of HCG 16

Although the quality of this first light EPIC observation is not all we might wish, and we are therefore unable to trace the diffuse emission from HCG 16 out to even as large a radius as was reached with ROSAT, the superior spectral resolution and sensitivity of XMM, compared to the ROSAT PSPC, allow us to establish some important facts about the group:

- There is substantial, genuinely diffuse X-ray emission in the centre of the group, surrounding the galaxies and extending to a radius of at least 135 kpc from the centre of the group.
- The mean spectrum of this emission is well represented by a thermal plasma with a low temperature (0.49 keV) and non-zero metallicity.
- The total extent and luminosity of this hot gas suggests that it cannot be attributed to galactic wind emission from the active galaxies within the group, and the morphology of the gas appears more relaxed than was apparent from the ROSAT data.
- The properties of this diffuse X-ray emission are subject to large errors, but seem to fit rather well onto trends established by other X-ray bright groups. HCG 16 appears to lie at the extreme poor end of the spectrum of X-ray emitting groups.

These properties suggest that HCG 16 is a genuinely bound group, with a potential well in which gas has collected and been heated, in a manner analogous to other X-ray bright groups. In other words, the group has collapsed, even if we cannot state if it has already virialized. A system which is dense in three-dimensions is also strongly suggested by the extraordinary degree of activity seen in its galaxies, which has presumably been triggered by galaxy interactions.

However, HCG 16 is so far unique amongst groups with detectable hot intergalactic gas, in containing no bright early-type galaxies. One possible explanation, is that HCG 16 is a rare example of a group which has collapsed in a single event (rather than building through a sequence of mergers) very recently, so that the processes which lead to the conversion of

late-type into early-type galaxies (mergers, interactions, stripping etc.) have not yet had time to run their course. If this picture is correct, then the implications are substantial. Typical X-ray bright groups, dominated by a central bright early-type galaxy, represent only a minority of groups. It is therefore very important to establish whether more typical spiral-dominated groups have similar gas content. If they do, then such groups may constitute the dominant component of the baryonic content of the Universe (Fukugita et al. 1998). The properties of HCG 16 lend tentative support to this idea.

There is one fly in the ointment. If HCG 16 is a system currently near maximum collapse, then its velocity dispersion should be rather *high* (Mamon 1993). However, after correction for measurement errors, the velocity dispersion of the group is actually unusually *low*: 99 km s^{-1} if only the central four galaxies catalogued by Hickson are considered, and 76 km s^{-1} if the additional three Ribeiro et al. (1996) galaxies are included (Mamon, private communication). In terms of the parameter $\beta_{\text{spec}} = (\mu m_p \sigma^2)/kT$, the ratio between the specific energy on the galaxies, and that in the gas, which is normally approximately unity in clusters, and rather lower in groups (Helsdon & Ponman 2000a), HCG 16 has the remarkably low value, $\beta_{\text{spec}} \sim 0.1$. This is difficult to understand, unless we happen to be viewing HCG 16 at an angle such that most of the galaxy motions lie within the plane of the sky.

In order to establish whether HCG 16 is truly a recently collapsed system with a fortuitous alignment of galaxy orbits, or whether we must seek some other explanation for its remarkable combination of properties, we require deep X-ray studies of other groups with similar optical characteristics. Such investigations should be carried out with XMM and Chandra over the next few years.

Acknowledgements. We are grateful to the referee, Gary Mamon, for interesting conversations on the status of HCG 16 and for the very useful comments of his referee report. We are grateful to Monique Arnaud for helpful discussions throughout this work and invaluable assistance in the estimation of the PSF. We also thank G.W. Pratt for contributions to the analysis and for very useful comments on the original manuscript. Our thanks to R. Gastaud and D.M. Neumann for software support and to R. Rothenflug for help in the initial stages of the analysis. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France. The paper is based on observations obtained with XMM-Newton, an ESA science mission with instruments and contributions directly funded by ESA Member States and the USA (NASA). EPIC was developed by the EPIC Consortium led by the Principal Investigator, Dr. M. J. L. Turner. The consortium comprises the following Institutes: University of Leicester, University of Birmingham, (UK); CEA/Saclay, IAS Orsay, CESR Toulouse, (France); IAAP Tuebingen, MPE Garching, (Germany); IFC Milan, ITESRE Bologna, IAUP Palermo, Italy. EPIC is funded by: PPARC, CEA, CNES, DLR and ASI.

References

- Arnaud, M., & Evrard, A. 1999, MNRAS, 305, 631
- Arnaud, M., Neumann, D. M., Aghanim, N., et al. 2001, A&A, 365, L80
- Babul, A., Balogh, M. L., Geraint, F. L., & Gregory, B. P. 2002, MNRAS, 330, 329

- Bahcall, N. A., Harris, D. E., & Rood, H. J. 1984, *ApJ*, 284, L29
- Belsole, E., Sauvageot, J. L., Arnaud, M., Ponman, T. J., & Rothenflug, R. 2001, Proceedings of the conference: Galaxy Clusters and the High Redshift Universe Observed in X-Rays, XXIth Moriond Astrophysics Meeting, ed. J. Trần Thanh Vân, & D. M. Neumann, in http://www-dapnia.cea.fr/Conferences/Morion_astro_2001/abs03/belsole.html
- Balogh, M. L., Babul, A., & Patton, D. R. 1999, *MNRAS*, 307, 463
- Bourdin, H., Slezak, E., Bijaoui, A., & Arnaud, M. 2001, Proceedings of the conference: Galaxy Clusters and the High Redshift Universe Observed in X-Rays, XXIth Moriond Astrophysics Meeting, ed. J. Trần Thanh Vân, & D. M. Neumann, in http://www-dapnia.cea.fr/Conferences/Morion_astro_2001/abs03/bourdin.html [astro-ph/0106138]
- Buote, D. A. 2000, *MNRAS*, 311, 176
- Buote, D. A., & Fabian, A. C. 1998, *MNRAS*, 296, 977
- Cavaliere, A., Menci, N., & Tozzi, P. 1997, *ApJ*, 484, L21
- Davis, D. S., Mulchaey, J. S., & Mushotzky, R. F. 1999, *ApJ*, 511, 34
- Dos Santos, S., & Mamon, G. A. 1999, *A&A*, 352, 1 (DSM99)
- Dos Santos, S., & Doré, O. 2002, *A&A*, 383, 450
- Evrard, A. E., Metzler, C. A., & Navarro, J. F. 1996, *ApJ*, 496, 494
- Fabbiano, G. 1988, *ApJ*, 330, 672
- Fukugita, M., Hogan, C. J., & Peebles, P. J. E. 1998, *ApJ*, 503, 518
- Finoguenov, A., & Ponman, T. J. 1999, *MNRAS*, 305, 325
- Ghizzardi, S. 2001, EPIC-MCT-TN-011 (XMM-SOC-CAL-TN-0022) <http://xmm.vilspa.esa.es/docs/documents/CAL-T-0022-1-0.ps.gz>
- Helsdon, S. F., & Ponman, T. J. 2000a, *MNRAS*, 315, 365 (Paper A)
- Helsdon, S. F., & Ponman, T. J. 2000b, *MNRAS*, 319, 933 (Paper B)
- Hernquist, L., Katz, N., & Weinberg, D. H. 1995, *ApJ*, 442, 57
- Hickson, P. 1982, *ApJ*, 255, 382
- Hickson, P., & Rood, H. J. 1988, *ApJ*, 331, L69
- Hickson, P., Mendes de Oliveira, C., Huchra, J. P., & Palumbo, G. 1992, *ApJ*, 399, 353
- Hickson, P. 1997, *ARA&A*, 35, 357
- Holland, A. D., Turner, M. J. L., Abbey, A. F., & Pool, P. 1996, *SPIE*, 2808, 414
- Lloyd-Davies, E. L. D., Ponman, T. J., & Cannon, D. B. 2000, *MNRAS*, 315, 689
- Mamon, G. A. 1986, *ApJ*, 307, 426
- Mamon, G. A. 1993, in Gravitational Dynamics and the N-Body Problem, ed. F. Combes, & E. Athanassoula, mtg. held in Aussois, France, March 1993 [astro-ph/9308032]
- Mulchaey, J. S., Davis, D. S., Mushotzky, R. F., & Burstein, D. 1996, *ApJ*, 456, 80
- Mulchaey, J. S., & Zabludoff, A. I. 1998, *ApJ*, 496, 73
- Mulchaey, J. S. 2000, *ARA&A*, 38, 289
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1995, *MNRAS*, 275, 720
- Pildis, R. A., Bregman, J. N., & Evrard, A. E. 1995, *ApJ*, 433, 514
- Ponman, T. J., Bourner, P. D. J., Ebeling, H., & Böhringer, H. 1996, *MNRAS*, 283, 690 (PBEB)
- Ponman, T. J., Cannon, D. B., & Navarro, J. F. 1999, *Nature*, 397, 135
- Riberio, A.L.B., de Carvalho, R.R., Coziol, R., Capelato, H.V., & Zepf, S.E. 1996, *ApJ*, 463, L5
- Saracco, P., & Ciliegi, P. 1995, *A&A*, 227, 301
- Stark, A. A., Gammie, C. F., Wilson, R. W., et al. 1992, *ApJS*, 79, 77
- Strüder, L., Briel, U., Dennerl, K., et al. 2001, *A&A*, 365, L18
- Tozzi, P., Scharf, C., & Norman, C. 2000, *ApJ*, 542, 106
- Tully, R. B. 1987, *ApJ*, 321, 280
- Turner, M. J. L., Reeves, J. N., Ponman, T.J., et al. 2001a, *A&A*, 365, L110 (Paper A)
- Turner, M. J. L., Abbey, A., Arnaud, M., et al. 2001b, *A&A*, 365, L27 (Paper B)
- Vikhlinin, A., Forman, W., & Jones, C. 1999, *ApJ*, 525, 47
- Walke, D. G., & Mamon, G. A. 1989, *A&A*, 225, 291
- Xue, Y. J., & Wu, X. P. 2000, *ApJ*, 538, 65

An *XMM-Newton* observation of the dynamically active binary cluster A1750

E. Belsole^{1,2}, G. W. Pratt¹, J.-L. Sauvageot¹, and H. Bourdin³

¹ Service d'Astrophysique, CEA Saclay, L'Orme des Merisiers, Bât. 709, 91191 Gif-sur-Yvette Cedex, France
e-mail: gwp@discovery.saclay.cea.fr; jsauvageot@cea.fr

² H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK

³ Observatoire de la Côte d'Azur, BP 4229, 06304 Nice Cedex 4, France
e-mail: bourdin@obs-nice.fr

Received 28 August 2003 / Accepted 19 November 2003

Abstract. We present results from the *XMM-Newton* observation of the binary cluster A1750 at $z = 0.086$. We have performed a detailed study of the surface brightness, temperature and entropy distribution and confirm that the two main clusters of the system (A1750 N and A1750 C) have just started to interact. From the temperature distribution, we calculate that they are likely to merge sometime in the next 1 Gyr. The more massive cluster, A1750 C, displays a more complicated temperature structure than expected. We detect a hot region associated with a density jump ~ 450 kpc east of the cluster centre, which appears to be a shock wave. This shock is not connected with the binary merger, but it is *intrinsic* to A1750 C itself. From simple physical arguments and comparison with numerical simulations, we argue that this shock is related to a merging event that A1750 C has suffered in the past 1–2 Gyr. The larger scale structure around A1750 suggests that the system belongs to a rich supercluster, which would presumably increase the likelihood of merger events. These new *XMM-Newton* data thus show that A1750 is a complex system, where two clusters are starting to interact before having re-established equilibrium after a previous merger. This merger within a merger indicates that the present day morphology of clusters may depend not only on on-going interactions or the last major merging event, but also on the more ancient merger history, especially in dense environments.

Key words. X-rays: galaxies: clusters – X-rays: individual: A1750 – X-rays: general – cosmology: large-scale structure of Universe – shock waves

1. Introduction

Clusters of galaxies are the largest gravitationally bound structures in the Universe, and because of their observable properties and sheer number, they represent invaluable cosmological probes.

The hierarchical scenario of structure formation in the Universe predicts the growth of clusters of galaxies through accretion of smaller units by gravitational infall and mergers. During a merger event more than 10^{63} ergs are dissipated in the intra-cluster medium (ICM) by shock heating, leading to strong variations of physical characteristics such as the temperature, density and entropy. Numerical N-body simulations have shown that mergers produce substructure which is detected in both the density and temperature distributions (Schindler & Müller 1993; Roettiger et al. 1997; Ricker 1998). The temperature substructures survive for ~ 4 – 6 times longer than density substructures, thus the temperature distribution is a strong indicator of the cluster history and dynamical state. X-ray

spectro-imaging observations are well suited to investigate in a deeper detail the observable characteristics of the merger event.

Density substructure in X-ray observations of some low-redshift clusters (Forman & Jones 1982; Mohr et al. 1993), together with the detection of other substructure in optical observations (Escalera et al. 1994), suggests that these clusters are dynamically young (Buote & Tsai 1996)

The first quantitative X-ray temperature maps were obtained with ASCA (e.g., Markevitch et al. 1998) but, as with the later BeppoSax observations (e.g., De Grandi & Molendi 1999), resolution was low due to the energy dependent Point Spread Function (PSF) of these telescopes. Thanks to the high effective area and spatial resolution now available from *XMM-Newton* and Chandra, precise spatially resolved temperature maps are now possible. Results confirm the increase in temperature in the merging regions (e.g., Arnaud et al. 2001; Markevitch & Vihkinen 2001; Markevitch et al. 2002; Neumann et al. 2003) but the new observations have also detected new phenomena, such as the cold fronts (Markevitch et al. 2000).

Send offprint requests to: E. Belsole,
e-mail: e.belsole@bristol.ac.uk

Substructures in clusters contain a fossil record of the merger history. Statistical studies of cluster morphology can thus provide an important test of cosmological models of structure formation. In practise, attempts to compare the prediction of numerical simulations of structure formation to quantitative statistical studies of nearby cluster morphology have been hampered by our current poor knowledge of the effect of cluster growth on morphology. Despite the large amount of data available, the dynamical evolution of the ICM and the relation between the galaxies and gas during a merger event is still poorly understood (see Buote 2002 for a review). A better understanding of the physics of merger events, in particular the relaxation time of substructures (Nakamura 1995) is required.

Hence a systematic study of merging clusters of galaxies would help a great deal in deepening our understanding of the process of cluster formation and evolution. Our *XMM-Newton* Guaranteed Time (GT) time program was established for this purpose. We selected a small sample of clusters displaying the signature of substructure (on the basis of ROSAT imaging analysis), and which could be considered as being in different *epochs* of a merger event.

In this paper we present the first object in our sample: the binary cluster A1750. In optical, A1750 shows a multi-peaked galaxy distribution (Beers et al. 1991; Ramirez & Quintana 1990; Donnelly et al. 2001); the two major peaks have a radial velocity difference of $\sim 1300 \text{ km s}^{-1}$. A1750 was observed in X-ray with *Einstein* (Forman et al. 1981), and subsequently identified as a canonical binary cluster, though three peaks are clearly visible in the *Einstein* X-ray image. The cluster was later observed with ROSAT and ASCA (Novicki et al. 1998; Donnelly et al. 2001); three peaks are also visible in the 2nd ROSAT image. The latter authors combine ROSAT and ASCA observations, as well as the galaxy distribution, giving the first kinematic and dynamical description of the cluster. They detected an enhancement of order 30% in the temperature between the two main peaks in the X-ray emission, and suggested that a shock region is developing in the gas by compression. However, Donnelly et al. (2001) concluded that the two clusters could effectively be considered as isolated objects, as the hot gas in the interacting region did not have a strong impact in the global temperature estimate, and its contribution to the total density was weak. Donnelly et al.'s mass analysis indicated a mass ratio of order 1.3, but their dynamical analysis was not conclusive as to whether or not the system is bound.

In this paper we use the high sensitivity and spatial resolution of *XMM-Newton* to gain new insights into this system. Taking advantage of the large field of view of *XMM-Newton*, we observed the two main subclusters in a single pointing. The exceptional sensitivity of *XMM-Newton* allows us to produce temperature and entropy maps with an accuracy which considerably surpasses previous attempts. We detect new features in all maps and relate these to the dynamical state of the system, arguing that, while the clusters are just beginning to interact, at least one of them (A1750 C), and possibly both, may not yet have relaxed from a previous merger event.

Throughout the paper we assume $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $q = 0.5$. In this cosmology 1 arcmin corresponds to 130 kpc at the mean redshift of the cluster ($z = 0.086$).

2. Observations and data preparation

2.1. Observations

A1750 was observed for 34 ks in July 2001 (Revolution 300) by *XMM-Newton*. In this paper only data from the European Photon Imaging Camera (EPIC; Strüder et al. 2001; Turner et al. 2001) are considered. Calibrated event list files were provided by the *XMM-Newton* SOC. The observations were obtained with the *MEDIUM* filter; the full frame mode was used for MOS and the extended full frame for pn. Throughout this analysis single pixel events for the pn data (PATTERN 0) were selected, while for the MOS data sets the PATTERNs 0–12 were used.

The background estimates were obtained using a blank-sky observation consisting of several high-latitude pointings with sources removed (Lumb et al. 2002). The blank-sky background events were selected using the same selection criteria (such as PATTERN, FLAG, etc.) used for the observation events. Furthermore, the blank-sky background file was recast in order to have the same sky coordinates as A1750, ensuring that the source and background products come from the same region of the detector, reducing errors induced by any detector position dependence.

The source and background events were corrected for vignetting using the weighted vignetting method described in Arnaud et al. (2001). This allows us to use the on-axis response matrices and effective areas.

2.2. Background estimate

The estimate of the background level is a crucial point since we are interested in extended and low surface-brightness sources, where the background effects are important especially at large distance from the centre.

The *XMM-Newton* background consists of several components, which may be variable in time and in space distribution. The soft proton background is a time and flux variable component. Within some observations it is possible to find several periods of time where the mean flux level varies by a factor of order 10 or more. To clean the data for this emission we reject all frames outside 2σ of the mean value in the 10–12 keV (12–14 keV for pn) light curves for each camera, using the method of Pratt & Arnaud (2002) to determine the mean and σ values. For this analysis, light curves were grouped in bins of 100 s. The observation was relatively clean: using these criteria $\sim 10\%$ of the total exposure time was rejected. The final exposure times are 30.4 ks, 31.4 ks, 23.4 ks for MOS1, MOS2 and pn cameras respectively. The blank-sky background event files were then cleaned using the same criteria.

A second component of the *XMM-Newton* background is represented by the particle induced background, which dominates at high energy ($>5 \text{ keV}$) and induces fluorescence lines (Al, Si, Cu, Au) from the shielding of the camera and the detector itself. This background component is effectively described by the blank-sky background, under the hypothesis that the variation of this particle background is small. It has been demonstrated that the particle background is variable at

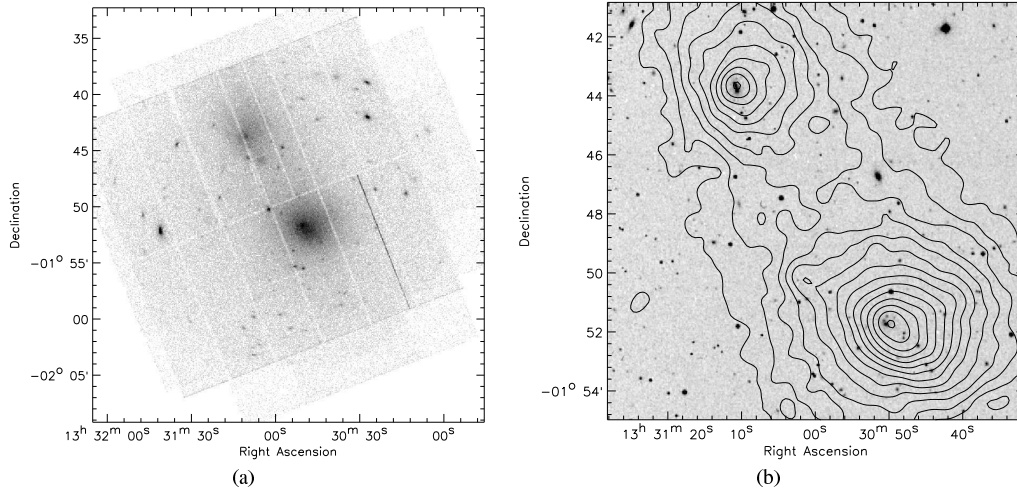


Fig. 1. a) EPIC counts image in the energy band 0.3–7.0 keV; the image is not corrected for vignetting **b)** Iso-intensity contours of the MOS1+MOS2 image in the energy band [0.3–1.4] keV overlaid on the optical DSS image. Contours are logarithmic, and were obtained by generating contours after replacing point sources with Poisson noise from surrounding annuli.

the $\sim 10\%$ -level. This variation is sufficiently small that it can be taken into account by a normalisation in the [10–12] keV (MOS) and [12–14] keV band (pn) between the source and the blank-sky background observations. The normalisation used in this work are 1.09, 1.11, 1.11 for MOS1, MOS2 and pn respectively.

Finally, we have to consider the astrophysical X-ray background, which is the combination of a soft ($E < 1.5$ keV) component, mainly due to the local bubble, and a hard component due to unresolved cosmological sources (mainly AGN). This emission is properly an X-ray component (i.e., not particle-induced, or due to soft proton flares). The hard component of this astrophysical X-ray background is well taken into account by the subtraction of the blank-sky background. At lower energies, however, the astrophysical X-ray background is variable across the sky (e.g., Snowden et al. 1997). To account for this, one possibility is to use a local background as described in Pratt et al. (2001) and Arnaud et al. (2002). However, emission from A1750 fills the whole field of view (FoV) and there is no large region which can be considered free of cluster emission with enough confidence to be used as a local background. In this analysis, the blank-sky background represents the best estimate we can give of the background associated with the cluster. However, the examination of radial profiles in several different bands suggests that at large radii the emission remaining after the subtraction of the blank-sky background is essentially of order zero, giving us confidence in our use of the blank-sky background only.

3. Morphology

3.1. Image

Figure 1a shows the combined MOS1, MOS2 and pn image of A1750 in the energy band 0.3–7.0 keV. The image is in

counts, has not been corrected for exposure, and the regions outside the field of view (FoV) have not been masked. A total of 356 000 photons have been collected by EPIC in this energy band. The double distribution of the emission is clear, with one emission peak at the centre of the FoV and the other to the north-east. There are also many serendipitous point sources and some extended sources: their positions and fluxes were obtained using the EMSRLI file in the pipeproducts, and all those with flux greater than 10^{-14} ergs s $^{-1}$ arcmin $^{-2}$ were masked throughout this analysis.

Contours of the low-energy (0.3–1.4 keV), adaptively-smoothed EPIC/MOS image are shown overlaid on the DSS image in Fig. 1b. The X-ray emission peaks are at $[\alpha = 13^{\text{h}}30^{\text{m}}49^{\text{s}}.881, \delta = -01^{\circ}51'46''.70]$ and $[\alpha = 13^{\text{h}}31^{\text{m}}10^{\text{s}}.941, \delta = -01^{\circ}43'41''.65]$ for the cluster at the centre of the FoV (hereafter A1750 C) and the north-eastern cluster (hereafter A1750 N) respectively.

The isophotes at the centre of A1750 C are elongated and their ellipticity decreases with the distance from the centre. The distance between the peak of X-ray emission of A1750 C and the cD galaxy visible in the DSS image is $\sim 15''$ (32.5 kpc). This shift is larger than the possible attitude error of *XMM-Newton*. The cD galaxy is offset from the X-ray peak towards the east, in which direction there is a clear compression of the isophotes.

The centroid of the X-ray emission of A1750 N is located exactly over the superimposed optical emission of two central galaxies. The X-ray isophotes in the central regions are quite circular but they show a slight compression to the east-south-east. The abrupt change in the shape and orientation of the isophotes at large scale is however remarkable, where they become very elliptical and oriented in the direction of the line joining the centres of the two clusters.

Table 1. Best fit results of the 2D β model.

	A1750 N	A1750 C
r_{c1} (kpc)	250	230
r_{c2} (kpc)	180	200
β	0.56	0.68
α_0 (J2000)	13 ^h 31 ^m 10 ^s .5	13 ^h 30 ^m 49 ^s .0
δ_0 (J2000)	-01°43'20".61	-01°51'54".14
PA	-12°12'53"	24°38'14"

3.2. 2D β model fitting

To take into account the binary nature of the cluster and to give a quantitative measure of the morphological characteristics we observe, and in an attempt to detect density substructures, we performed a bi-dimensional analysis.

We adopted the hypothesis that a β model (Cavaliere & Fusco-Femiano 1976) is an appropriate description of a relaxed isothermal cluster. We fit the surface brightness distribution of A1750 at low energy (where the density distribution is least temperature dependent) with a β model and quantify the deviation from this model and the presence of substructures. In fact, the interest of this exercise is to compare A1750 with the surface brightness distribution of two relaxed clusters.

For the fit we followed the prescription of Neumann & Böhringer (1997) and apply this analysis to the MOS camera only because the large gaps in the pn camera are relatively difficult to take into account. Images of pixel size of 4".1 from the weighted MOS event files and in the energy band 0.3–1.4 keV were created, and then summed. Since the weight correction is applied directly to the events, the statistical Poisson distribution of photons does not apply for these images. To correctly take into account the errors, the variance is given by $\sigma^2 = \sum_i w_i^2$ where w_i is the weight associated to the i th photon (see Arnaud et al. 2001). The error images were generated for each MOS image and the quadratic sum of the latter gives the summed error image. The combined MOS image was smoothed with a Gauss filter of $\sigma = 15''$, chosen to be of the same order of the MOS PSF. This smoothing allows us to be confident that in each spatial interval of integration (meta-pixel) we can use the Gaussian statistic and find the best χ^2 when fitting. The error image was treated according to the error propagation function for Gaussian filtered images described in Neumann & Böhringer (1997).

Because of the double nature of A1750, the 2D β model for a single cluster (see Neumann & Böhringer 1997; Pratt & Arnaud 2002) has been modified in order to take into account simultaneously the two clusters. The background was approximated as a unique constant value. The region between the two clusters was masked before fitting so as to reduce the interference from a possible interaction region. Point sources were also masked. Results of the best fit double 2D β -model are shown in Table 1. The best fit model was then subtracted from the image in order to quantify the significance of any possible excess flux. The significance was calculated following the Appendix of Neumann & Böhringer (1997). The positive residuals, traced at 2, 3, 5 and 7 σ above the background are shown in Fig. 2.

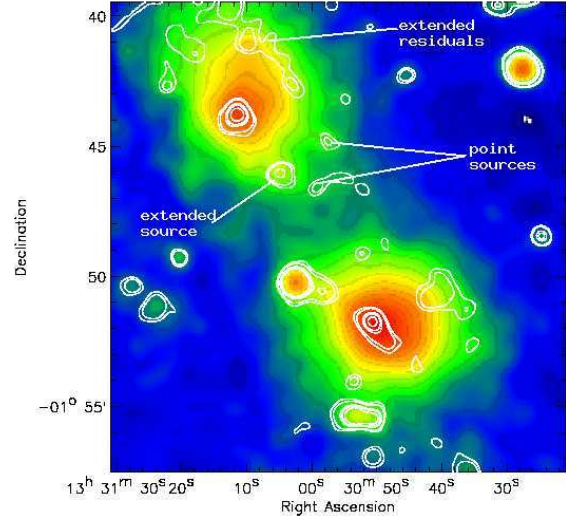


Fig. 2. EPIC/MOS adaptively smoothed image with, overlaid, contours of the residuals after the subtraction of a double β model. The residuals are traced at 2,3,5 and 7 σ above the background. Further details in the text. (See the electronic edition of the *Journal* for a color version of this figure.)

The image clearly shows significant residuals in the centre of each cluster. The core of A1750 N is off-centre with respect to the structure at larger scale (which is well described by the β -model). The same effect is marginally visible at the centre of A1750 C. Significant residuals, other than the two cores, are detected to the north of A1750 N and to the west of A1750 C.

This analysis does not show any significant substructure in the region between the two clusters (hereafter the middle region), with the exception of two marginally detected ($\sim 2\sigma$) point sources and a further source which is detected at greater than 5σ . From comparison with the raw image, it is clear that this source is extended.

We have considered the nature of the extended source in the middle region. It is clearly detected in emission in the pn image (see Fig. 1a), but unfortunately most of the emission is located in the gaps of the MOS images. We extracted the spectrum of this source in a circle of 45". As a background we used an annular region centred on the source, with an inner radius 1/3 and outer radius 2/2. The spectrum does not contain emission lines strong enough to estimate the source redshift. We fitted the spectrum with an absorbed MEKAL model with redshift and column density parameters fixed to those of A1750 ($z = 0.086$ and $N_H = 2.37 \times 10^{20} \text{ cm}^{-2}$). The best fit gives $kT = 4.6^{+2.3}_{-1.3} \text{ keV}$ and a flux of $1.3 \times 10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1}$ in the energy range [0.3–10.0] keV. A superposition of the residual contours on the DSS image does not show any obvious optical counterpart.

4. Thermal structure

The lack of significant substructure in the middle region seems to suggest that there is little or no interaction between the two

Table 2. Results of the global spectrum analysis in the region shown in Fig. 3. The column density is obtained by fitting the MOS spectra only and then fixed to this value when combining with the pn spectrum. The value of the column density listed here is the best fit MOS value (see text for more details). Errors are at 90% for one significant parameter.

region	N_{H} (10^{20} at/cm $^{-2}$)	kT (keV)	Z solar	$\chi^2/\text{d.o.f}$	L_{x} [0.2–10] keV (erg s $^{-1}$)
Cluster C	2.18 ± 0.44	3.87 ± 0.10	0.32 ± 0.04	1217.79/1139	2.2×10^{44}
Cluster N	1.00 ± 0.8	2.84 ± 0.12	0.22 ± 0.05	797.33/663	1.3×10^{44}
Middle	$0.5(<1.4)$	$5.12^{+0.77}_{-0.69}$	$0.27^{+0.28}_{-0.19}$	418.32/426	1.7×10^{43}

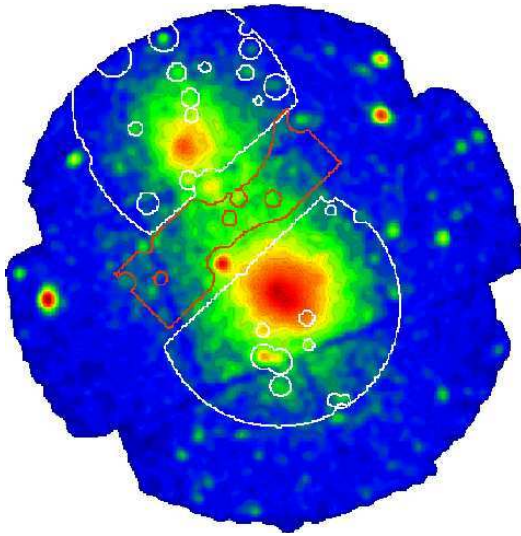


Fig. 3. EPIC/MOS adaptively smoothed image in the energy range 0.3–6.0 keV with superimposed contours of the 3 regions of interest for the global spectrum analysis. The image is logarithmically scaled. (See the electronic edition of the *Journal* for a color version of this figure.)

components of the cluster (however, see Sect. 7.1). In other words, the merger event is sufficiently early that the density enhancement expected in the interaction region (e.g. Roettiger et al. 1997; Ritchie & Thomas 2002) has not yet been produced (or is only just beginning). Knowledge of temperature structure is thus a further important step toward the comprehension of the dynamical state of the system.

The spectral analysis was performed by extracting spectra from the weighted source and background (blank-sky) event lists. The out of time events affecting the pn observation were not taken into account, but in our experience they are not expected to have a detectable effect on the results. The response files used in this analysis are: m1_mediumv9q20t5r6.all.15.rsp (MOS1), m2_mediumv9q20t5r6.all.15.rsp (MOS2) and epn_ef20_sY9_medium.rsp (pn).

4.1. Global spectrum

We obtained a global spectrum for each of the regions displayed in Fig. 3. The regions were chosen to give a first measure of the spectral characteristics of the two clusters and the middle region, and also serve as a useful comparison with previous results. The particular shape of the middle extraction region is determined by the exclusion of point sources and cluster emission. The spectrum of the background was extracted in the same region and, after normalisation, subtracted from the source spectrum.

The spectra of the 3 cameras, excluding the channels below 0.3 keV, were simultaneously fitted with an absorbed MEKAL model with the temperature, metallicity and normalisation (emission measure) as free parameters. The column density was fixed to the best fit value obtained by fitting the MOS cameras only. This value is in agreement with the galactic value for A1750 C. (The spectrum of A1750 N and the middle region are better fitted with a lower than Galactic absorption). The best fit results are listed in Table 2. The temperature in the middle region is significantly higher (by $\sim 30\%$) than that of A1750 C, in agreement with the results of Donnelly et al. (2001). In order to be confident in this detection of a significantly higher temperature, we also investigated the dependence of the temperature of middle region spectrum on the Galactic absorption. If we fix the N_{H} to the Galactic value, the temperature drops to 4.35 ± 0.45 keV (for a reduced $\chi^2 = 425.86/427$ d.o.f.), while if we fix the N_{H} to the value obtained by fitting cluster N, the best fit temperature is $4.9^{+0.6}_{-0.5}$ keV ($\chi^2 = 418.98/427$ d.o.f.). In both cases the temperature is higher than the temperature of A1750 C at the 90% confidence level, and the goodness of the fit is clearly in favour of a lower absorption. In Table 2 we list the temperatures obtained by fitting simultaneously the three cameras with the column density fixed to the best fit MOS value in each region. The spectral fits are shown in Fig. 4.

For A1750 C and the middle region, our results are in excellent agreement with the ASCA analysis (Donnelly et al. 2001), but we find a lower temperature for A1750 N, if the mean temperature between zones 1 and 2 in Donnelly et al.'s paper is taken. However, if we compare our result with the ASCA temperature in region 1 of Donnelly et al.'s Fig. 2, the two results are in better agreement.

We further fit the global spectrum of A1750 N with a two temperature model. In this case we find $kT_1 = 3.17 \pm 0.1$ keV, $kT_2 = 0.65 \pm 0.17$, with abundances of 0.3 ± 0.7 solar and a $\chi^2 = 754.05/660$ d.o.f. According to the F -test, it is 99%

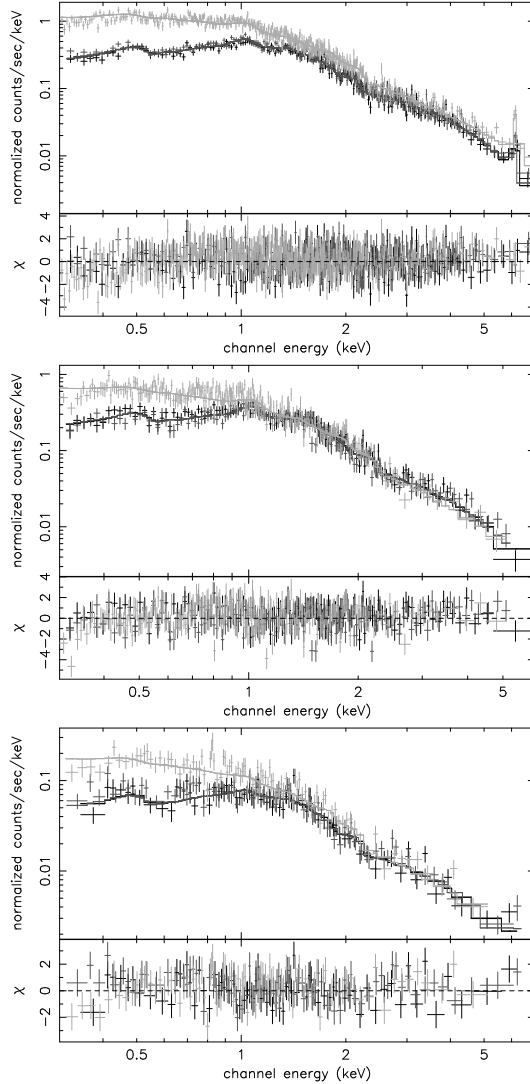


Fig. 4. Global spectra of the three regions in Fig. 3: A1750 C (Top); A1750 N (centre); middle region (bottom). MOS spectra are in black, the pn in grey. The three camera spectra were fitted simultaneously.

probable that the 2 temperature model is a better representation of the data, suggesting that the gas is multi-temperature (see also Sect. 6).

4.2. Temperature map

4.2.1. Adaptive binning

The high sensitivity of *XMM-Newton* allows us to build a temperature map. To optimise the spatial resolution, without any spatial a priori, we have defined a specific binning which follows the statistics, and leads to fairly homogeneous errors in the

temperature determination for each meta-pixel. This approach is similar to the one presented by Sanders & Fabian (2001). We produce a full energy band image with sources excised. The algorithm starts from the largest spatial scale and continues down to the smallest possible, dividing by 4 each cell of the image containing more than a given number of photons (here we choose 2000 photons for the 3 cameras in the full energy band prior to background subtraction). We end up with a list of cells that have more or less the same number of photons. We apply this spatial binning to the observation and to the background to extract all spectra. Then using XSPEC, we fit the MEKAL model (with the N_H fixed to the galactic value and abundances set to 0.3 solar) to each spectrum. The fit temperature of each cell is then used to build the temperature map shown in Fig. 5a. The typical error (which depends slightly on the temperature) is around ± 0.6 keV.

As already suggested by the broad global spectrum analysis above, the region between the two clusters displays a higher temperature, implying a dynamical interaction. The map is in excellent agreement with that of Donnelly et al. (2001), but because of the improved sensitivity, our map indicates that the temperature structure is more complicated.

A1750 N displays a relatively uniform temperature with a suggestion of a cooler core region (although the large bins are not optimised). To the west side of A1750 N we observe a higher temperature which may be connected with the interaction between the two clusters. A hot region to the south of the centre of A1750 N should be also noted, which interestingly lies ahead of the compression of the isophotes near the core.

A new feature, observed for the first time in these data, is the hot region ~ 3 –4 arcmin to the east of the centre of A1750 C. This hot region corresponds exactly to a compression of the isophotes in the X-ray surface brightness, and does not appear to be related to the larger hot region due to the interaction between the clusters. In other words, it appears to be *intrinsic* to A1750 C itself.

4.2.2. Wavelet temperature analysis

We also computed a temperature map by applying the new multi-scale spectro-imagery algorithm described in Bourdin et al. (2004). In that paper, the algorithm was tested on simulated observations; here we present preliminary results from the first application to a real observation. For a detailed description of the method see Bourdin et al. (2004).

The A1750 temperature map was computed using only the event lists from the MOS cameras, as we do not yet have a stable model for the pn background. Due to the steep decrease of the cluster emissivity with radius, the analysis has been restricted to regions where significant signal is detected at more than one scale. The background normalisation was fixed to a value estimated for each camera in an external region of the field of view, where the background emissivity dominates. The temperature map is a 128×128 pixels image, where the multi-scale analysis has been performed on 5 scales in order to prevent large scale edge effects, so that details on spatial scales from around $15''$ to $4'$ are expected. Spectral fits with

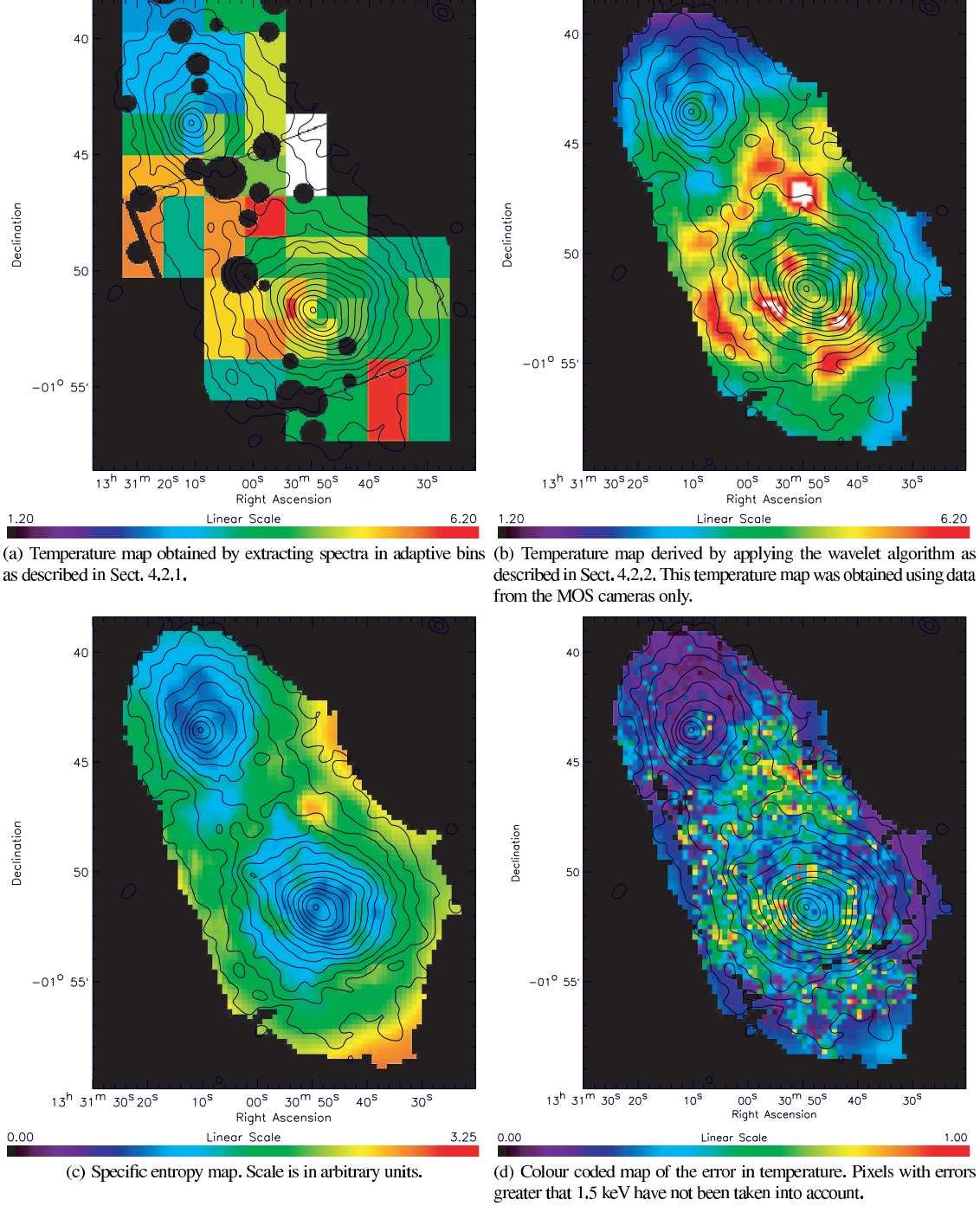


Fig. 5. Temperature and entropy maps of A1750. Temperature colour scales are in keV.

error-bars higher than 1.5 keV are not taken into account due to evidence for a bias of the temperature estimator in regions with high background level. The issue of background modelling for spectral fits in each element is the major source of uncertainty, especially at large distance from the centre where the background itself begins to dominate. To be conservative, we have applied a posteriori a mask to the wavelet maps defined such that at least 10 photons are detected in the emissivity wavelet map. This corresponds to a 3σ signal above the background.

Since we have not yet perfected the removal of point sources in the algorithm, the following wavelet-derived maps are presented without point sources removed. The side by side comparison of the adaptive binning and wavelet temperature maps (Figs. 5a and 5b) allows the interested reader to judge the effect of the point sources on the temperature structure.

Note that, in the wavelet map, all the point sources appear in the highest frequency plane. This means that their contribution to the local temperature is limited to that spatial frequency (in other words they always appear in the temperature map at the highest frequency, i.e., as point-like sources).

In general, the wavelet-derived temperature (Fig. 5b; errors Fig. 5d) map is in good agreement with the adaptively-binned temperature map. We note the following:

- The hot rectangle to the South of A1750 C in the adaptively binned map is not seen in the wavelet map, but this region is at the periphery of the cluster emission and has large errors ($kT = 5.7^{+1.9}_{-1.7}$ keV).
- The hot regions appearing in the wavelet map to the south and south-west of the centre of A1750 C appear to be due to point sources.
- The hot region between the clusters is recovered in the wavelet map, but is compromised by a point source. Because of this, we are unable to determine if the structure is continuous or clumpy.
- The temperature structure the east of the centre of A1750 C (discussed in more detail later on) is robust, as there are no associated point sources. It is resolved in the wavelet map into two arc-like structures of higher temperature. The first is located at ~ 1.8 from the centre of A1750 C to the east. The second one starts at around 3.5 in the same direction. These two hot regions are separated by a gap of lower temperature.
- Finally, the hot region at the Northwest end of the arch between the clusters appears to be stable, appearing in both maps, although some of this emission may be due to a strong point source.

The western side of A1750 C appears to be relatively isothermal. A1750 N displays a quite uniform temperature distribution, with an indication of a higher temperature in the centre than that observed with the adaptive binning map (but see also Sect. 6).

Note that the wavelet temperature map has been derived using only the MOS data, while the adaptive binning map uses data from all three cameras. As such, we would expect the wavelet map to only have approximately half the sensitivity of the adaptively-binned map. This loss of sensitivity is crucial

in regions of low surface brightness, such as the middle region. Despite these caveats, the agreement between the maps is good.

5. Entropy

The multi-scale spectro-imagery algorithm described above can be applied to estimate the projected spatial distribution of other parameters of the intra-cluster gas, provided that the expected parameter and its fluctuation are estimated within each resolution element. In this paper, we applied the algorithm to the EPIC-MOS data to compute a crude map of the gas specific entropy, which is convenient to define in terms of observed quantities as $S \propto T/n^{2/3}$, as a function of the intra-cluster gas temperature T , and density n ¹.

To do so the averaged cluster emissivity per surface unit (N) and temperature (T) are fitted simultaneously within the different resolution elements, following the same procedure as for computing the temperature map. Since the cluster emissivity (N) scales as to the square of the density of the intra-cluster gas (n), both quantities leads to a crude estimator of the averaged specific entropy, $\hat{S} = \frac{T}{N^{1/3}}$.

Just like for the temperature map, the “entropy” map is obtained by computing maps of $\hat{S} \pm \sigma_{\hat{S}}$ at five different scales, leading to a wavelet transform. Then the wavelet coefficients are thresholded in order to remove the noise contribution, again following a 1σ significance criterion.

The entropy map, shown in Fig. 5c is very similar to entropy maps obtained in numerical simulations of cluster mergers (e.g. Ricker & Sarazin 2001). We observe a lower entropy in the centres of the two clusters, with an increase towards the external regions. The entropy distribution of A1750 N is very elongated towards the north, mainly following the emissivity distribution. We observe that the eastern side of A1750 C displays an enhancement in the entropy which does not follow a spherical distribution. The projected entropy distribution is quite turbulent in this region. The entropy in the middle region is also higher than in the centre of the two clusters, as expected from shock heating.

6. One dimensional analysis

The two dimensional analysis above has revealed a wealth of new detail in both the morphology and temperature of the gas in this cluster. In this Section we use the two dimensional results to divide each subcluster into discrete regions, with the aim of undertaking a classical 1D analysis for comparison with previous results. The regions we have chosen are shown in Fig. 6. These regions are:

- Region 1 (R1) corresponds to the part of A1750 N which is relaxed-looking and fairly isothermal. R1 has been defined by excluding all data between position angles 77° – 330° (anticlockwise from N , centred on $\alpha = 13^{\text{h}}31^{\text{m}}00^{\text{s}}$; $\delta = -01^\circ46'46''$), related to all hot regions connected with the interaction.

¹ More precisely, the definition of “entropy” is the logarithm of the quantity above plus a constant (see Ponman et al. 2003 and references therein).

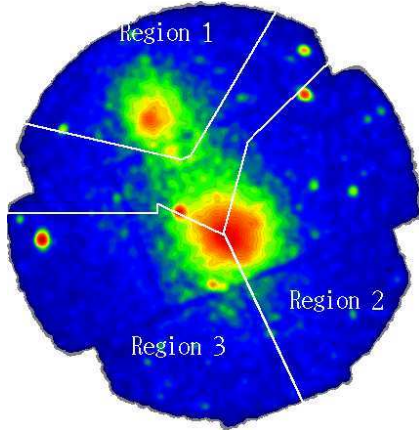


Fig. 6. Adaptively-smoothed MOS image in the energy band 0.3–6.0 keV. The lines limit the regions used to extract 1 dimensional profiles. The image is in logarithmic scale. (See the electronic edition of the Journal for a color version of this figure.)

- Region 2 (R2) is that part of A1750 C which is relaxed-looking and very isothermal. R2 is defined by a sector between position angles 205° – 354° , anticlockwise, centred on $\alpha = 13^h30^m49^s.881$, $\delta = -01^\circ51'46''.70$, the A1750 C emission peak.
- Finally, Region 3 (R3) is delimited by a sector between position angles 65° – 205° (anticlockwise), centred on the A1750 C emission peak. R3 includes the zone $\sim 4'$ the east of the centre which shows the hot, arc-like temperature feature and isophotal compression.

6.1. Data analysis

6.1.1. Surface brightness profiles

In each region, we extracted azimuthally averaged surface brightness profiles from the observation and the corresponding Region in the blank-sky background. Photons were binned in circular sector annuli of width $5''$ centred on the peak of the emission. We use the energy band [0.3–1.4] keV to minimise the dependence of the emissivity on the temperature; the upper energy limit was set to avoid the instrumental Al and Si lines².

The background surface brightness profiles were subtracted from each corresponding cluster profile, using the appropriate [10.–12.]/([12.–14.]) MOS(pn) keV normalisation factor. Each camera was treated separately. In all cases, the MOS and pn profiles agree well barring the expected normalisation differences, and so in each case the MOS and pn profiles were co-added.

Except in one case (R3), the surface brightness profile was binned such that (i) at least a S/N ratio of 3σ was reached,

² We checked the emissivity-temperature dependence of the FeL blend by extracting profiles with the [0.9–1.2] keV band excluded. The best fitting β -model parameters did not change significantly in this case, thus to maximise the S/N , we use the entire [0.3–1.4] keV band.

and (ii) the width of the bins increased with radius, with $\Delta(\theta) > 0.1\theta$. The logarithmic binning ensures a roughly constant S/N in the outer parts of the profile.

6.1.2. Temperature profiles

Radial temperature profiles were produced for each region using the same mask as was used for the corresponding surface brightness profiles. Spectra were extracted in circular sector annuli centred on the peak of the emission. The widths of the annuli were chosen so that a minimum 5σ detection was achieved in the [2.0–5.0] keV band. A minimum width of $30''$ was also imposed, corresponding to the 90% encircled energy radius of the MOS PSF. The blank-sky background spectra were extracted from the same regions and subtracted from the source spectra. The background subtracted spectra were binned to 3σ to allow the use of Gaussian statistics.

Spectra from the three cameras were fitted simultaneously with an absorbed MEKAL model, with the column density fixed to the best-fit of the corresponding global spectrum (Table 2). The abundance for the fits was frozen at the best-fit global value for whichever cluster was under consideration³. The MOS spectra were fitted in the range [0.3–10.0] keV, and the pn spectra in the range [0.4–10.0] keV.

We stress that we have not attempted to undertake a full PSF and deprojection analysis and that the results presented here are *projected* temperature profiles. PSF effects will be important if the cluster possesses very significant temperature gradients, or if the surface brightness profile is very peaked. We have taken care to minimise PSF effects by ensuring that the bins have a minimum width greater than the 90% encircled energy radius of the PSF, and we have already seen that the temperature gradients, while significant, are relatively mild (not nearly so strong as in a cooling flow, for example).

The temperature plots below are presented with the upper X-axis in arc-minutes. The lower X-axis shows the radius in terms of the virial radius, calculated using the average temperature from the global spectral fit (Sect. 4.1), and the $r_{200} - T$ relation of Evrard et al. (1996) at the cluster redshift, viz:

$$r_{200} = 3690 (T/10 \text{ keV})^{1/2} (1+z)^{-3/2}. \quad (1)$$

6.1.3. Gas density profiles

The surface brightness profile of each region was fitted with various parametric models, all of which were convolved with the *XMM-Newton* PSF (Ghizzardi 2001; Griffiths & Saxton 2002) and binned in the same way as the observed profile. The surface brightness profile, at low energy, is generally well represented by a β -model:

$$S(\theta) = S_0 \left(1 + \frac{\theta^2}{\theta_c^2} \right)^{-3\beta+0.5} \quad (2)$$

³ We did fit the radial profiles with spectra where the abundance was a free parameter: at 90% confidence level however, the temperature profiles do not show significant differences from those obtained with the abundance frozen at the global value.

Table 3. Results of the BB model analytical fits to the gas surface brightness profiles, errors are 90% confidence.

Parameter	R1	R2
	A1750 N	A1750 C W
$n_{H,0} (h_{30}^{1/2} \text{ cm}^{-3})$	7.68×10^{-3}	5.89×10^{-3}
$r_c (') / \text{kpc}$	$1.96^{+1.83}_{-0.17} / 254.4$	$2.03^{+0.15}_{-0.12} / 263.5$
β	$0.51^{+0.05}_{-0.04}$	$0.69^{+0.03}_{-0.02}$
$R_{\text{cut}} (') / \text{kpc}$	$2.04^{+2.64}_{-0.23} / 264.8$	$1.05^{+0.21}_{-0.16} / 136.3$
$r_{\text{c,in}} (') / \text{kpc}$	$0.12^{+0.04}_{-0.04} / 15.6$	$0.10^{+0.02}_{-0.02} / 13.0$
$\chi^2/\text{d.o.f.}$	57.4/35	42.2/42

^a The maximum value of $r_{\text{c,in}}$ is fixed to $1'$.

where S_0 is the central intensity, θ_c the core radius and β the slope. In the hypothesis of an isothermal cluster, the two parameters θ_c and β are related to the gas density profile by:

$$n_{\text{gas}}(r) = n_0 \left(1 + \frac{r^2}{r_c^2} \right)^{-3\beta/2}. \quad (3)$$

In addition to the standard β -model above, we have also used the BB parametric model (Pratt & Arnaud 2002), which is a double isothermal β -model which assumes that both inner and outer gas density profiles can each be described by a β -model, but with different parameters. The boundary between the two regions is a free parameter of the fit, and the density distribution, and its gradient, are continuous across the boundary.

6.2. A1750 N; R1

A single β -model as described above, while being a relatively good description of the outer regions, is not a good description of the entire profile ($\chi^2 = 122.5$ for 37 degrees of freedom (d.o.f.)). As is often found in these cases, progressively excluding the central regions improves the fit. The best fit β -model, obtained by excluding the inner $\sim 1'$ yields $\chi^2 = 48.9$ for 27 d.o.f.; in this case $\beta = 0.46$ and $r_c = 1.4'$. We thus used the double isothermal BB model to fit the profile. This resulted in a much better fit to the *entire* radial range: $\chi^2 = 57.4$ for 35 d.o.f. The BB model fit is shown in Fig. 7; best-fit results are shown in Table 3.

The best-fit outer β value of the BB model fit is in excellent agreement with the 2D analysis in Sect. 3.2 (results in Table 1), and with the ROSAT-derived results presented by Donnelly et al. (2001). The fact that this is so, when we are not fitting exactly the same regions, is an indication that the results derived here are very robust.

The radial temperature profile is shown in Fig. 8; note that this is the first such profile for this cluster. The cluster is detected with excellent S/N up to $\sim 0.4r_{200}$, the limit being due to the edge of the *XMM-Newton* FoV. The temperature profile in this Region shows a surprising amount of variation. However, while the variations are significant in terms of their errors, they are actually quite small. For instance, the lowest temperature annular sector has $kT = 2.6$ keV, and the highest $kT = 3.3$ keV, an absolute variation of 0.7 keV.

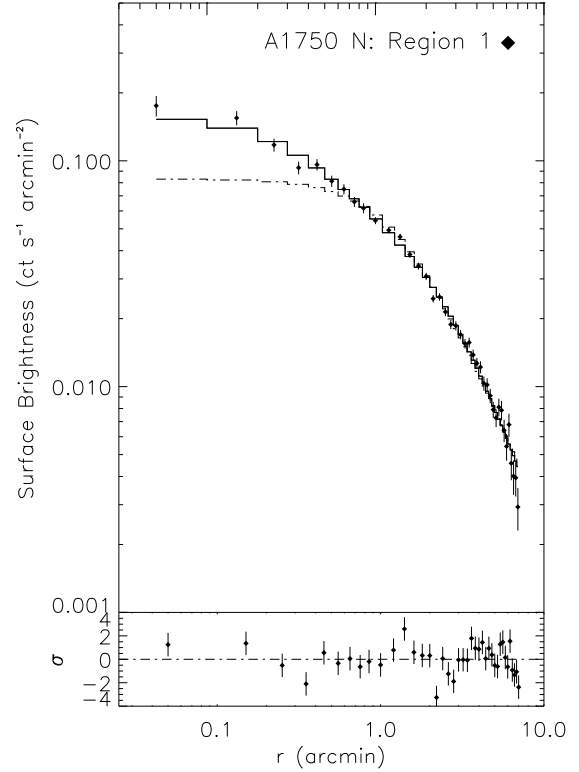


Fig. 7. The surface brightness profile of A1750 N, Region 1, binned as described in the text. It has been background subtracted and corrected for vignetting. The profile is shown with the best fit BB model; the residuals are shown in the bottom panel. The dash-dotted line shows the best-fit single β -model to the outer regions.

There is a dip towards the centre, which may be due to a cooling flow (discussed further below in Sect. 7.3). From the second to the final annular sector, there is a smooth, gentle decline broken only by the fifth annular sector. However, the general features of this temperature profile are in good agreement with the spatially-resolved temperature map (Sect. 4.2).

6.3. A1750 C; R2

R2 extends to the eastern side of A1750 C, and corresponds to the part of the cluster which shows the least amount of structure in isophotes and appears quite isothermal in the temperature map.

Once again, a single β -model fails fully to describe the data ($\chi^2 = 56.7$ for 44 d.o.f.), although the central excess is far less evident than for R1. The best fitting single β -model requires the exclusion only of the central $0.7'$, and yields $\chi^2 = 36.6$ for 37 d.o.f. A BB model with parameters given in Table 3 is an excellent fit to the entire profile ($\chi^2 = 42.2/42$ d.o.f.). The profile, together with the best fitting BB model, is shown in the right hand top panel of Fig. 9.

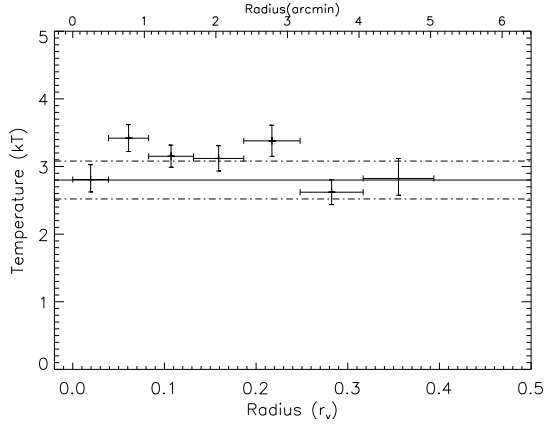


Fig. 8. Temperature profile of A1750 N, Region 1. The straight line is the best fit global spectrum temperature, the dash-dotted lines denote $\pm 10\%$ of this value. Errors are at 1σ for one interesting parameter.

The best-fit outer β value of the BB model fit is in good agreement with the 2D analysis in Sect. 3.2, and with the ROSAT-derived results presented by Donnelly et al. (2001). Once again, this gives us good confidence in our results.

The temperature profile for R2 is shown in the right hand bottom panel of Fig. 9; this is the first temperature profile for this cluster.

The cluster is detected with excellent S/N up to $\sim 0.6r_{200}$. In this Region, as expected from the temperature map, the temperature profile is quite smooth, and shows little variability.

6.4. A1750 C; R3

Here we come to the region of A1750 C which displays isophotal compression and, compared to R2, considerable temperature structure.

The surface brightness profile of R3 is shown compared to the profile from R2 in the left-hand top panel of Fig. 9; in order to avoid smoothing of the discontinuity, we have not adopted logarithmic binning. The difference between the profiles is striking: the R3 profile is more peaked, as is shown by the overplotted best-fitting BB model to R2. In addition, there appears to be a change of slope between 3 and 4 arcmin from the centre, the shape of which may indicate a discontinuity in the gas density profile.

For simplicity, we assume spherical symmetry. To quantify the discontinuity, we fitted the surface brightness profile of R3 with a radial density model composed of a β -model and a power law separated by a jump:

$$n(r) = \begin{cases} n_2(r) = A_2 \left(1 + \frac{r^2}{r_c^2}\right)^{-3\beta/2} & \text{if } r < r_{\text{cut}} \\ n_1(r) = A_1 \left(\frac{r}{r_{\text{cut}}}\right)^\alpha & \text{if } r > r_{\text{cut}} \end{cases} \quad (4)$$

where

$$n_1|_{r_{\text{cut}}} = X_{\text{jump}} \times n_2|_{r_{\text{cut}}}.$$

In this model, n_2 and n_1 are the gas densities before and after the discontinuity, respectively. The free parameters of the fit were the slope β and coreradius of the β -model, the slope α of the power law, and the position (r_{cut}) and amplitude X_{jump} of the density jump. The profile was fitted between bin 2 ($10''$) and bin 60 ($6'$). The best fit density model is shown both alone and superimposed on the data in Fig. 10. The goodness of the fit gives an excellent $\chi^2 = 51.4/53$ d.o.f. The best fit parameters are: $\beta = 0.41$, $r_c = 0.6$, $\alpha = 0.85$. The jump position is $r_{\text{cut}} = 3.44'$ and the amplitude of the density jump is $X_{\text{jump}} = 0.78^{+0.09}_{-0.1}$, i.e., a 20% jump in density.

We also changed the bin range within which we fit the model and also used a second model composed of 2 power laws. In both cases we obtain consistent results within the errors, including jump amplitudes, indicating that the results are robust.

The orientation of the discontinuity with respect to the line of sight is unknown: projection effects may act to reduce the apparent density jump at the discontinuity. If we do not see the discontinuity exactly perpendicular to the line of sight, the sharpness of the edge will be reduced. In other words, it is likely that our derived value for the density jump is a lower limit.

The radial temperature profile for R3 is shown in the bottom left panel of Fig. 9. Here, the ring edges were chosen in order to match the density jump described above, and to have at least a 5σ signal above the background in the $[2.0-5.0]$ keV band.

The temperature profile for this region is quite variable. Even bearing in mind that both temperature profiles in Fig. 9 are plotted with a logarithmic x-axis, it is obvious that the profile of R3 is more variable than that of R2. The maximum of the temperature profile at ~ 1.5 (in fact a 1σ deviation) corresponds to the peak of the inner high-temperature arc seen in the temperature map (Fig. 5b). This temperature peak does not correspond to any significant structure in the surface brightness profile, even if the surface brightness profile at this distance from the centre does display some level of discontinuity. It is worth noting that within the errors we do not detect any temperature discontinuity at $3.4'$, the position of the jump in the density profile. We further note that there is a significant temperature jump between annuli 5 and 6. However this does not correspond to any significant jump in the observed surface brightness profile.

6.5. Entropy profiles

With a view to making a comparison with relaxed clusters, we determined the entropy profiles of Regions R1, R2 and R3. The entropy was determined from the BB analytical model fit to each gas density profile and the observed (projected) temperature profiles, taking, as is now customary in this field, $S = T/n_c^{2/3}$. For R3, we use the gas density model described above in Sect. 6.4, limiting the radial values to the fit limit of $r < 6'$. The resulting entropy profiles are shown, plotted in terms of the virial radius, in Fig. 11. Typical errors, corresponding to the error in each bin of the temperature profile, are

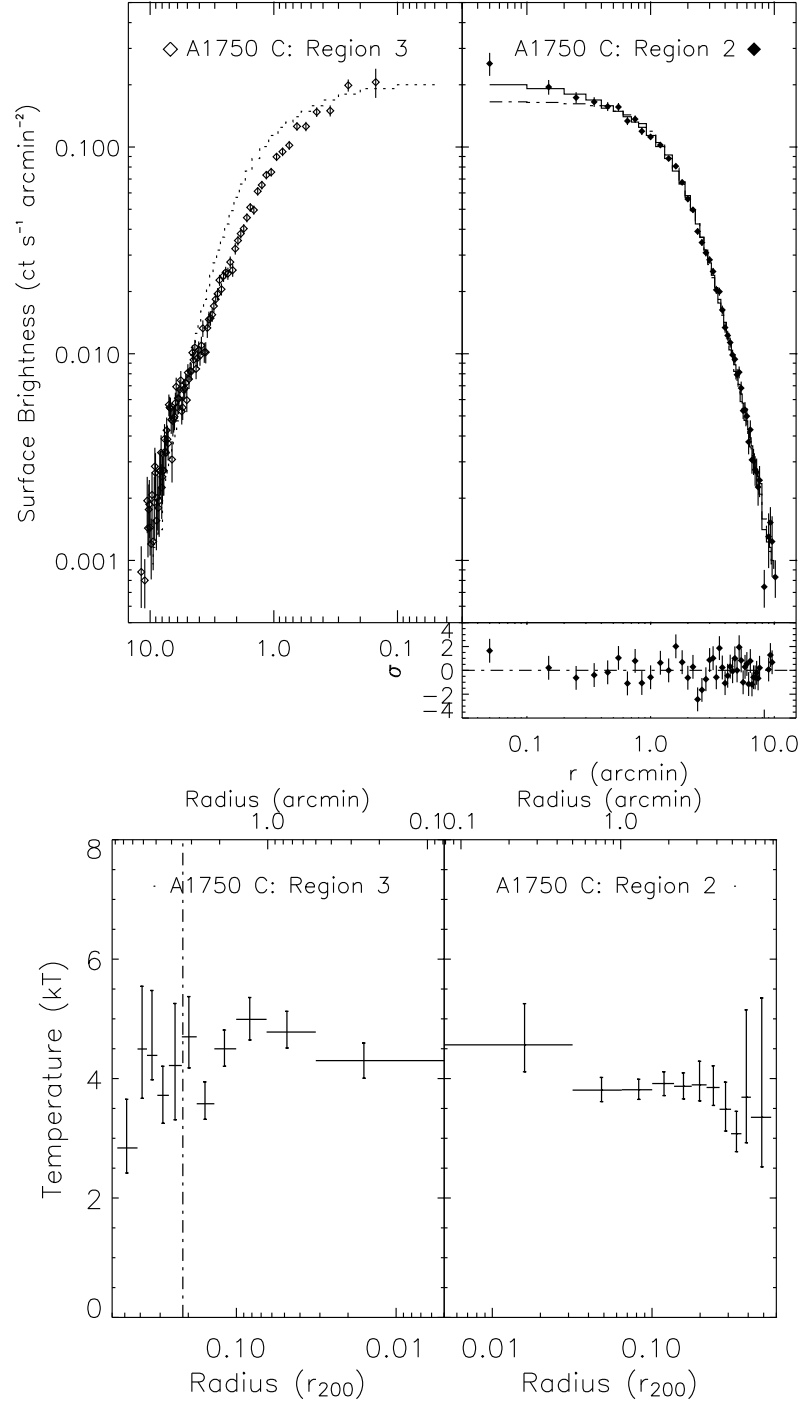


Fig. 9. The surface brightness (top) and temperature (bottom) profiles of Regions 2 and 3 of A1750 C. The right hand panel in each case shows R2, the left hand panel shows R3. For R2, the dot-dash line is the best-fit β -model and the solid line is the best-fit BB model. The residuals are also shown. The dotted line overplotted on R3 is the best-fit BB model for R2.

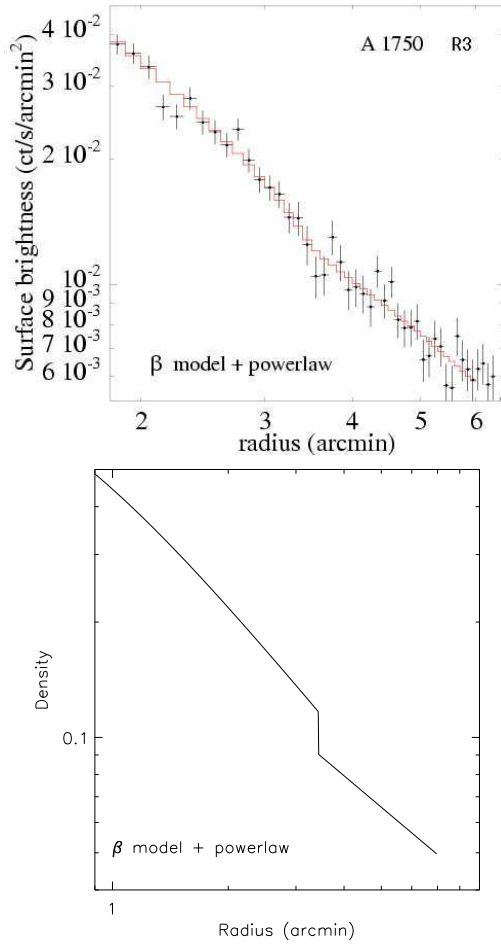


Fig. 10. *Top:* surface brightness profile of A1750 C, R3: data points and density model used to fit the data. *Bottom:* model used to fit the surface brightness profile (see Eq. (4)). The model is normalised to 1 at the centre.

indicated. The overplotted line shows the $S \propto r^{1.1}$ behaviour expected from shock heating (Tozzi & Norman 2001).

7. Discussion

7.1. The current merging event

In this section we discuss a possible scenario for the current merging event (i.e., that between A1750 N and A1750 C) based on simple assumptions.

The weak enhancement in temperature between the two clusters (Fig. 5a) is a strong indication that the clusters are just beginning to interact.

We can attempt to calculate the Mach number and the collision velocity from the observed temperature variations. In this analysis, we follow the simple assumptions outlined in Markevitch et al. (1999), viz: that the merger is symmetric

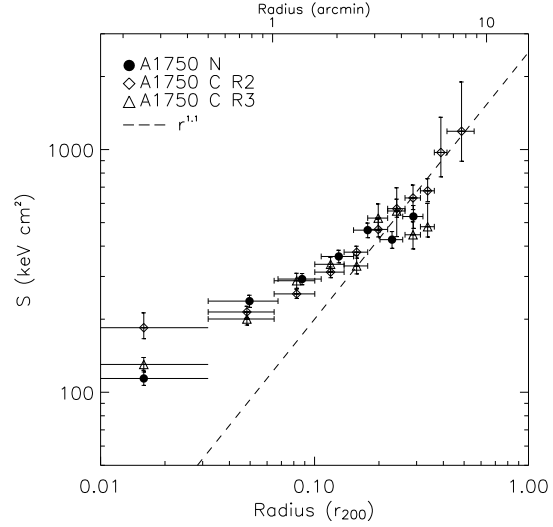


Fig. 11. The entropy profiles of Regions R1, R2 and R3. The dashed line corresponds to $S \propto r^{1.1}$, with an arbitrary normalisation. The profile for R3 has been limited to the range of our fit to the surface brightness ($r < 6'$).

in the plane of the sky, that the shocked gas is on average at rest with respect to the centre of mass, and that the shocked gas is nearly isothermal and in equipartition. The Mach number and collision velocity can then be obtained by applying the Rankine–Hugoniot jump conditions for a one-dimensional shock, following Landau & Lifshitz (1959; see also Markevitch et al. 1999; Sarazin 2002):

$$\frac{1}{C} = \left[4 \left(\frac{T_2}{T_1} - 1 \right)^2 + \frac{T_2}{T_1} \right]^{1/2} - 2 \left(\frac{T_2}{T_1} - 1 \right); \quad (5)$$

$$\mathcal{M}^2 = \frac{3C}{4 - C}; \quad (6)$$

where Eq. (5) corresponds to Eq. (2) of Markevitch et al. (1999) when $\gamma = 5/3$; T_1, T_2 are the pre-shock and post-shock temperatures and $C = n_2/n_1$ is the shock compression.

For this rough estimate, we assume that the pre-shock temperature ($kT = 3.10 \pm 0.14$ keV) can be approximated by the average temperature on the sides of each cluster symmetrically opposite the shock region (see Markevitch et al. 1999), and the post-shock temperature ($kT = 5.12 \pm 0.73$ keV) is given by the observed global temperature of the middle region (Sect. 4.1). We obtain $C = 1.89$ and a Mach number of $\mathcal{M} = 1.64$.

Assuming that the surface brightness scales as the square of the density, using the above results we expect an increase of ~ 3.6 in the surface brightness between the clusters. This is not a huge effect. To compare directly with the observations we integrated the surface brightness perpendicular to the axis joining the cluster centres, and compared this to the same integration of an image of the best-fitting BB models. We then computed the ratio in brightness between the middle region and the region symmetrically opposite the clusters (in a similar

fashion to the determination of the Mach number above), for both the image and the model⁴. The range of ratios we found (1.5–6.2) brackets the expected increase.

To first order, the time until first core passage (t_{merge}) between A1750 C and A1750 N can be obtained by considering the distance of the two components and the distribution of the temperature and the density, even if projection effects are the major limitation in the accuracy of such an analysis.

The centres of A1750 C and A1750 N lie at a projected distance of $675 h_{50}^{-1}$ kpc. If we assume that the difference in radial velocity is exclusively due to the infall of A1750 N onto A1750 C (in the reference frame of A1750 C), the two systems would be so close – well within the virial radius of each cluster – that we should observe far larger density/surface brightness distortions and temperature enhancements in the middle region (as seen in e.g., Fig. 3 of Ricker & Sarazin 2001). On the other hand, if we assume that the physical distance is determined by the Hubble flow, meaning that the motion of the two clusters is exclusively radial, the difference in their mean radial velocities would yield a distance between the two clusters of $\sim 20 h_{50}^{-1}$ Mpc. This corresponds to nearly 10 times the virial radius of either cluster and thus the two should not display any sign of interaction whatsoever, in disagreement with the results described in detail above. The real distance surely lies in between these two values, and judging by the observed characteristics, is likely to be less than the virial radius of A1750 N (~ 1.7 Mpc), using the $r_{200} - T$ relation of Evrard et al. (1996 – Eq. (1)).

Using Eq. (1) of Markevitch et al. (1999), we calculate a collision velocity of 1400 km s^{-1} , which is roughly of the same order as the radial velocity difference between A1750 C and A1750 N. For a rough, order of magnitude estimate of t_{merge} , we first assume the distance between the clusters to be the projected distance. In this case, we find $t_{\text{merge}} = 0.47$ Gyr. If we now assume that the distance between the clusters is given by the estimated virial radius of A1750 N, we find $t_{\text{merge}} = 1.3$ Gyr.

We can also compare our results with simulations of merger events of systems of nearly equal mass (Roettiger et al. 1996, 1997; Ricker & Sarazin 2001; Ritchie & Thomas 2002; Teyssier 2002), which resemble very closely the observed characteristics. We thus expect the two clusters to lie at a real distance close to, but somewhat lower than, their virial radii, around $1 h_{50}^{-1}$ Mpc. This would imply that the clusters will reach core passage some time in the next ~ 0.7 Gyr.

7.2. A1750 C: An unrelaxed cluster

Four pieces of evidence lead us to conclude that A1750 C is not a relaxed cluster, and further, that the perturbed state is intrinsic to the cluster itself and is not connected to the nascent merger with A1750 N. These are:

1. The observed discontinuity in the gas density profile.

A shock front is the most natural interpretation for the density discontinuity observed in A1750 C, R3 (Sect. 6.4).

In addition, the entropy map and profile of R3 (Fig. 11) show exactly the type of behaviour expected from a shock (the denser side has higher entropy).

Heating shocks are expected on theoretical grounds when clusters collide. Furthermore, such shocks are expected to be relatively weak, with a distribution of Mach numbers peaked at $M = 1.4$ (Gabici 2003). Shocks were proposed to explain the peculiar temperature and density structures observed in Cygnus A and A3667 (Markevitch et al. 1999) with ASCA and ROSAT. A clear detection of a shock wave was obtained with Chandra in the 1E 0657-56 cluster (Markevitch et al. 2002). Other weaker shocks were observed in A3667 (Vikhlinin et al. 2001) and A85 (Kempner et al. 2003).

By applying the Rankine–Hugoniot jump conditions (Eq. (6)) we can obtain the Mach number of the shock in A1750 C with respect to the ambient gas (assuming again an adiabatic index for monoatomic gas $5/3$). These conditions are valid if the gas within the shocked region is nearly isothermal and equipartition between electrons and ions applies. From the density jump in Sect. 6.4 we obtain $M = 1.19^{+0.13}_{-0.09}$. The expected temperature ratio for a shock with such a Mach number is $T_2/T_1 = 1.18$. Comparing the pre-shock $T_1 = 4.28^{+0.71}_{-0.62}$ keV and post-shock $T_2 = 4.59^{+0.56}_{-0.42}$ keV temperatures (see Sect. 6.4), we obtain a ratio $T_2/T_1 = 1.07(\pm 0.21)^5$, and thus, within the errors, the two independent measures are consistent.

From the jump conditions we can also infer the pressure before and after the shock, obtaining a pressure jump of a factor 1.5. This demonstrates that the gas in cluster C is not in pressure balance and thus hydrostatic equilibrium conditions should not be applied.

The geometry of the shock suggests that it is related to a merging event that A1750 C has suffered in the past. The merger velocity is $v_{\text{merge}} \sim 500 \text{ km s}^{-1}$, and thus taking the distance from the centre, a rough estimate of the time of the A1750 C merger gives $t_{\text{merge}} \sim 0.9$ Gyr. By comparing with simulations, the thermo-dynamical status this corresponds to an advanced phase of merger, at least 2 Gyr after core passage (Ricker & Sarazin 2001), which also corresponds to a sound crossing time⁶. Given our approximations, possible projection effects and the physical changes in the medium during the merger, and the difficulty of comparing with simulations, we cannot really be more precise. The observed temperature jump between bins 5 and 6 in the R3 temperature profile brings to mind a cold front. However, the propagation of a cold core into a hotter medium should also produce a density jump, as observed in e.g., A2142 (Markevitch et al. 2000). We have not been able to find such a discontinuity at this position, even by using profiles extracted in elliptical rings. However, given possible projection effects, we cannot entirely rule out the possibility of a cold front. Obviously, the better spatial

⁵ Assuming symmetrical errors for T_1 and T_2 .

⁶ $t_s = 6.6 \times 10^8 (T/10^8 \text{ K})^{-1/2} (D/\text{Mpc})$ where D is the cluster diameter (see Sarazin 1986 Eq. (5.54)).

⁴ The definition of the outer regions is somewhat arbitrary since we do not know the real distance between the clusters.

resolution of a Chandra observation would enable us to provide a definitive answer to this question.

2. **The elliptical isophotes at small scale, the offset of the core from the expected position given the larger scale structure, and the shift of the X-ray centroid from the position of the brightest cluster galaxy (BCG).**

These are particularly evident after subtraction of a 2D β -model (Fig. 2), and in Fig. 1b.

In their simulations of merging clusters Roettiger et al. (1997) predict a late phase of the interaction, after core passage, when the dark matter and the gas separate and the gas sloshes in a potential which is starting to re-establish a new equilibrium configuration. This phase is long in the timeframe of the merger event (at least up to 5 Gyr in the Roettiger et al. simulations). This sloshing of the gas can produce an elongated distribution as observed in A1750 C as well as a centroid shift (see also Roettiger et al. 1996). A similar gas distribution is observed in A1795 (Markevitch et al. 2001; Fabian et al. 2001; Ettori et al. 2002). In this case, however, the extent of the sloshing gas is 1/2 of that of A1750 C and is more closely connected with the central galaxy and the cooling flow. Note that Markevitch et al. (2001) and Ettori et al. (2002) concluded that the centre of A1795 is not in hydrostatic equilibrium even at larger scale. This is thus further strong observational evidence that A1750 C is in the late phase of an old merger event.

3. **The lack of evidence for a cooling flow.**

The flat central temperature gradient (Fig. 9) argues quite strongly against the presence of a cooling flow. We confirm this by computing the cooling time using Eq. (7). This gives $t_{\text{cool}} = 1.05 \times 10^{10}$ yr, of the same order as the Hubble time at the redshift of the cluster. Simulations of merging clusters suggest that only nearly head-on major mergers between objects of nearly the same mass can destroy cooling flows (e.g., Ritchie & Thomas 2002). Thus the lack of cool gas in the centre of A1750 C is likely the result of an old – and rather violent – merger event.

4. **The entropy profile, which appears to differ significantly from other high quality *XMM-Newton* profiles of relaxed clusters.**

We show in Fig. 12 the entropy profiles of R2 and R3, compared with the profiles of the relaxed clusters A1983 ($kT = 2.1$ keV; Pratt & Arnaud 2003) and A1413 ($kT = 6.9$ keV; Pratt & Arnaud 2002). The profiles have been scaled with the empirically-derived scaling relation ($S \propto T_X^{-0.65}$) of Ponman et al. (2003). While both R2 and R3 appear approximately to converge to the $S \propto r^{1.1}$ value expected from shock heating (Tozzi & Norman 2001) at a radius of $\sim 0.2r_{200}$, each exhibits a very high scaled central entropy value for its temperature (or mass), compared to the scaled entropy of the relaxed clusters.

In form, the entropy profiles of R2 and R3 are qualitatively similar to those found in the merger simulations of Ricker & Sarazin (2001). The entropy map is also reminiscent of their simulations. The increase in core entropy due to a merger has also been seen in the simulations of Ritchie & Thomas (2002). The increase appears to be the result of the mixing of shocked, high entropy gas from the outer regions

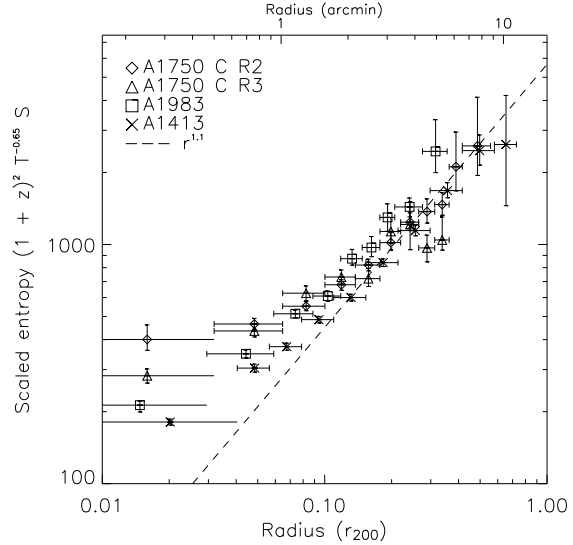


Fig. 12. The scaled entropy profiles of Regions R2 and R3 compared with the entropy profiles of A1413 and A1983 (Pratt & Arnaud 2002, 2003). The profiles have been scaled with using the empirically-derived relation $S \propto T_X^{-0.65}$ of Ponman et al. (2003). The dashed line corresponds to $S \propto r^{1.1}$, with an arbitrary normalisation.

into the low entropy core of the cluster. Thus the entropy profile is another strong piece of evidence that A1750 C has not yet relaxed from a previous merger event.

Crucially, all of the evidence listed above solely concerns A1750 C. These data strongly suggest that the cluster is in the late phase of a merger for which the core passage happened 1–2 Gyr ago. The cluster is in the process of re-establishing equilibrium, but this is likely to be interrupted by the new merging event occurring with A1750 N.

7.3. A1750 N

The morphological analysis of A1750 N (Sect. 3.2) has shown the elongated shape of the cluster, as well as an excess above a standard β -model which appears considerably off-centre with respect to the overall cluster morphology.

In discussing the form of this cluster we should first mention that at this off-axis angle PSF effects are expected to become important. In order to have a qualitative idea of this effect, we have simulated the whole observation in SCISIM, using as input the best-fitting 1D β -model values listed in Table 3. The resulting image confirms that emission from A1750 N should indeed be smeared by the PSF, preferentially in the azimuthal direction.

With this in mind, we now turn to the ROSAT image of the cluster, shown in Fig. 13, and discussed in more detail below (Sect. 7.4). It is evident that A1750 N seems elongated in the NE direction; however, given the larger field of view of ROSAT, we notice that there is another extended source just to the North of the cluster. We can thus speculate that there is

Table 4. Multi-temperature and CF fits of the inner annulus of A1750 N. The F -test is computed against the fit for a single temperature absorbed MEKAL model. Errors are 90% confidence.

Parameter	1T	2T	CF
kT_1 (keV)	$2.77^{+0.45}_{-0.35}$	$3.01^{+0.79}_{-0.41}$	$3.28^{+1.4}_{-0.46}$
kT_2 (keV)	–	$0.73^{+1.47}_{-0.23}$	–
Z/Z_\odot	$0.2^{+0.22}_{-0.16}$	$0.25^{+0.39}_{-0.15}$	$0.21^{+0.21}_{-0.14}$
\dot{M} ($M_\odot \text{ yr}^{-1}$)	–	–	$4.97 (<9.6)$
$\chi^2/\text{d.o.f.}$	108.22/100	105.15/98	105.77/98
F_{prob}	–	75.6%	67.4%

The low temperature of the CF model T_{low} was 0.71 keV.

some contamination by emission from this source, and indeed the 2D β -model analysis of the *XMM-Newton* observation suggests extended residuals in the direction of this source. Again we caution that PSF effects play a role at this distance from the centre.

Finally, we cannot entirely rule out that A1750 N has been disturbed, or is being disturbed, by interactions. It is possible that the observed isophotal compression to the South of the centre (see Fig. 2) is due to the interaction with A1750 C (suggesting a non-zero impact parameter), but with the present observation we cannot definitively say whether there is interaction with the source to the North.

Donnelly et al. (2001) speculate on the presence of a cooling flow in A1750 N. This is qualitatively in agreement with what we observe: a temperature drop is observed in the core of A1750 N (Fig. 8), together with an excess above a standard β -model.

In order to verify the presence of a possible cooling flow in A1750 N, we calculated the cooling time using (Sarazin 1986):

$$t_{\text{cool}} = 2.9 \times 10^{10} \text{ yr} \left(\frac{kT_X}{1 \text{ keV}} \right)^{1/2} \left(\frac{n_e}{10^{-3}} \right)^{-1}. \quad (7)$$

Using the central gas density obtained in Sect. 6.2, $n_e(0) = 7.68 \times 10^{-3} \text{ cm}^{-3}$; Eq. (7) gives $t_{\text{cool}} = 6.4 \times 10^9 \text{ yr}$, about half of the age of the Universe at the redshift of the cluster. This is consistent with a weak cooling flow. We thus fitted the MOS spectrum of the central bin with more complicated models:

- The sum of two MEKAL models absorbed by a column density fixed to the global spectrum fit value. The abundances of the two components are tied together.
- The sum of a MEKAL model and a cooling flow model (MKCFLOW), with fixed absorption and abundances of the CF model tied to the thermal model. Also the upper temperature of the CF model is tied to the temperature of the MEKAL model.

The results are shown in Table 4 together with the simple MEKAL model results for this bin. There is not a significant probability that the 2T model or the MEKAL+CFLOW model are a better representation of the data. All fits are very good and the small mass deposition rate found by fitting the CF model is in agreement with a small cooling flow.

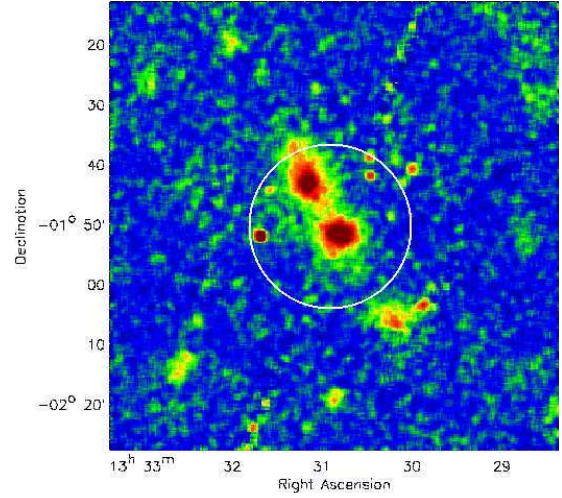


Fig. 13. ROSAT/PSPC image smoothed with a Gaussian filter of $\sigma = 20''$. The white circle traces the limit of the XMM field of view. (See the electronic edition of the Journal for a color version of this figure.)

Apart from the cooler core region, A1750 N displays a fairly uniform temperature distribution. With the present data, all evidence suggests that the only physical process which is disturbing A1750 N today is the merger with A1750 C. However, we stress that a conclusive analysis needs an observation of the whole cluster, and the data presented here are limited by the off-axis position in the *XMM-Newton* field of view (as discussed further below in Sect. 7.4).

7.4. At larger scale

The *XMM-Newton* FoV has allowed us to observe the two main clusters of A1750 in one pointing. However, other structures have been observed at larger scale with *Einstein* (Jones & Forman 1999). Figure 13 shows a smoothed ROSAT/PSPC image obtained from the ROSAT archive: the superimposed circle indicates the *XMM-Newton* FoV. To the south of A1750 C we observe another obviously extended source. This source has been observed in the optical and is discussed in Beers et al. (1991), where it is called A1750 S; it lies at a projected distance of $\sim 2 \text{ Mpc}$ from A1750 C and it is at the same redshift as A1750 C. Beers et al. (1991), using optical virial mass estimators, suggest that A1750 S contributes $\sim 2\%$ to the total mass of the system. It is worth noting that the gas distribution to the south-western side of A1750 C appears elongated in the direction of A1750 S.

Figure 13 also clearly indicates a further extended source just to the north of A1750 N, which is outside the FoV of our *XMM-Newton* observation. We recall that the subtraction of a 2D β -model (Sect. 3.2) suggested that there are further residual structures to the north of A1750 N, likely linked to this source. Other possibly extended sources are visible further to the north of A1750 N in the PSPC image. All these sources are aligned in

the direction of the line joining the centres of the 3 main clusters, strongly suggesting a filamentary structure. Einasto et al. (2001) found that A1750 belongs to a rich super cluster composed of 7 members. We searched clusters in the redshift range 0.08–0.09 in a $16^\circ \times 16^\circ$ field centred on the centre of this super cluster and found 13 clusters, including A1750 which is located at the north-eastern edge.

Given this observed large scale structure, it is not surprising that A1750 C has suffered one (or several) previous mergers along the putative filament and indeed, statistical studies suggest a correlation between the level of substructures and the environment density (Schuecker et al. 2001). This adds a further piece of evidence in favour of our interpretation. This may well be the case for A1750 N as well, where its elongated gas distribution may also be the result of accretion along the filament.

7.5. Caveat emptor

The scenario outlined above explains in a self-consistent manner the data we have analysed. In particular, the observed temperature and surface brightness structure would be difficult to explain if the clusters have already past one another (as suggested for A3528 by Gastaldello et al. 2003). We stress, however, that projection effects can be very important (e.g., as shown by Roettiger et al. 1997), and comparison with numerical simulations is necessarily a qualitative exercise. For a better understanding of this complex system, two more *XMM-Newton* pointings, one each centred on A1750 N and A1750 S, would be very useful, as would deep optical observations.

8. Summary and conclusions

We have reported the *XMM-Newton* GT observation of the merging cluster A1750. The main conclusions of this work may be summarised as follows:

- In the morphological analysis with 2D β -models, we detect excess and off-centre emission in the cores of both A1750 N and C. We do not find any significant substructure in the region between the two clusters after subtraction of the β -models. However, on close examination of the surface brightness in the region between the two clusters, we detect an increase of the same order as expected if the region between the clusters is a weak shock region (Sect. 7.1). We further detect excess residual emission to the north of A1750 N, which is likely (from comparison the *XMM-Newton* and PSPC images) related to the larger scale structure of the system.
- We have produced temperature maps by applying two different algorithms. We measured a temperature increase of order 30% in the region between A1750 N and C. A1750 N exhibits a relatively smooth, uniform temperature distribution, but there are significant temperature variations within A1750 C, which appear intrinsic to A1750 C and not connected to the merger between the two clusters.
- We use the 2D information to select interesting regions for a 1D analysis. For the regions corresponding to the relaxed, isothermal-looking parts of each cluster (R1 and R2),

the 1D and 2D β -model parameters are in excellent agreement with each other and with previous ROSAT analysis, although the *XMM-Newton* data require an additional central component properly to fit the profiles. Both temperature profiles appear to decline modestly (<10%) with radius.

- The density profile of the disturbed side of A1750 C (R3) is better described by a β -model and a power law, separated by a jump of order 20%. The temperature profile of this region is very irregular, but appears constant across the jump.
- Compared to the scaled entropy profiles of relaxed clusters, the scaled entropy profiles of both regions of A1750 C exhibit a high scaled entropy value in the centre, similar to what is observed in merging cluster simulations. Only beyond $\sim 0.2r_{200}$ does the profile of R2 attain the $S \propto r^{1.1}$ expected from shock heating.
- In the assumption of an head-on merger between A1750 C and N we calculated the Mach number and velocity of the merger from the measured temperatures. With the resulting values ($M = 1.64$ and $v_{\text{merg}} \sim 1400 \text{ km s}^{-1}$), we estimate that the two clusters will reach core passage in – roughly – less than 1 Gyr.
- By applying the Rankine–Hugoniot jump condition to the measured density jump in A1750 C (R3) we confirm the detection of a weak shock of $M = 1.2$, in agreement with the measured temperature of the pre-shock and post-shock gas. Under simplified assumptions and by comparison with numerical simulations, we estimate that a merger – intrinsic to A1750 C – has occurred sometime in the past 1–2 Gyrs.
- An inspection of the larger scale X-ray image suggests that A1750 N and C are the two main clusters within a filament which includes at least another small cluster at the same redshift of A1750 C (A1750 S). At still larger scale, A1750 lies in a filamentary supercluster containing, in total, 13 clusters.

We thus confirm that A1750 is a merging cluster in an early phase, when the two units have just started to interact and stand at a distance slightly lower than their virial radii. However, the global dynamical status of the system is far more complicated than expected. We conclude that the main cluster (A1750 C) has already suffered a previous merger and is now in the phase of re-establishing equilibrium. This phase will be interrupted by the current merger with A1750 N. As far as we know, this is the first time such complex dynamical signatures have been observed in clusters at this merger phase.

The present day morphology of clusters may thus depend not only on on-going mergers (or the last merger) but also on the more ancient merging history, especially in dense environments. This has to be taken into account in the interpretation of statistical studies of cluster morphology.

Acknowledgements. We are very grateful to Dr. M. Arnaud for her participation in the scientific discussion and help in the analysis of the density profiles. We thank R. Teyssier for useful discussions. The authors thank the anonymous referee for the interesting comments which improved the manuscript. This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France. The paper is based on observations obtained with *XMM-Newton*, an ESA science mission

with instruments and contributions directly funded by ESA Member States and the USA (NASA). EPIC was developed by the EPIC Consortium led by the Principal Investigator, Dr. M. J. L. Turner. The consortium comprises the following Institutes: University of Leicester, University of Birmingham, (UK); CEA/Saclay, IAS Orsay, CESR Toulouse, (France); IAAP Tuebingen, MPE Garching, (Germany); IFC Milan, ITESRE Bologna, IAUP Palermo, (Italy). EPIC is funded by: PPARC, CEA, CNES, DLR and ASI.

References

- Amaud, M., & Evrard, A. 1999, MNRAS, 305, 631
- Amaud, M., Aghanim, N., Gastaud, R., et al. 2001, A&A, 365, L67
- Amaud, M., Neumann, D. M., Aghanim, N., et al. 2001, A&A, 365, L80
- Amaud, M., Majerowicz, S., Lumb, D., et al. 2002, A&A, 390, 27
- Beers, T. C., Gebhardt, K., Forman, W., Huchra, J., & Jones, C. 1991, AJ, 102, 1581
- Bourdin, H., Sauvageot, J. L., Slezak, E., Bijaoui, A., & Teyssier, R. 2004, A&A, 414, 429
- Buote, D. A., & Tsai, J. C. 1996, ApJ, 458, 27
- Buote, D. A. 2002, in *Merging Processes in Galaxy Clusters*, ed. L. Feretti, I. M. Gioia, & G. Giovannini. Astrophysics and Space Science Library, 272 (Dordrecht: Kluwer Academic Publishers), 79
- Cavaliere, A., & Fusco-Femiano, R. 1976, A&A, 49, 137
- Crone, M. M., Evrard, A. E., & Richstone, D. O. 1996, ApJ, 467, 489
- De Grandi, S., & Molendi, S. 1999, ApJ, 527, 25
- Donnelly, R. H., Forman, W., Jones, C., et al. 2001, ApJ, 562, 254
- Einasto, M., Einasto, J., Tago, E., Müller, V., & Andernach, H. 2001, AJ, 122, 2222
- Escalera, E., Biviano, A., Girardi, M., et al. 1994, ApJ, 423, 539
- Ettori, S., Fabian, A. C., Allen, S. W., & Johnstone, R. M. 2002, MNRAS, 331, 635
- Evrard, A. E., Metzler, C. A., & Navarro, J. F. 1996, ApJ, 469, 494
- Fabian, A. C., Sanders, J. S., Ettori, S., et al. 2001, MNRAS, 321, L33
- Forman, W., Bechtold, J., Blair, W., et al. 1981, ApJ, 243, L133
- Forman, W., & Jones, C. 1982, ARA&A, 20, 547
- Gabici, S. 2003, in *Proceeding of the National Meeting on High Energy Astrophysics*, Roma 15-16 May 2003. To appear in the Frascati Physics Serie [astro-ph/0307499]
- Gastaldello, F., Ettori, S., Molendi, S., et al. 2003, A&A, 411, 21
- Ghizzardi, S. 2001, EPIC-MCT-TN-011 (XMM-SOC-CAL-TN-0022) <http://xmm.vilspa.esa.es/docs/documents/CAL-T-0022-1-0.ps.gz>
- Jones, C., & Forman, W. R. 1999, ApJ, 511, 65
- Kempner, J., Sarazin, C. L., & Ricker, P. M. 2002, ApJ, 579, 236
- Kempner, J., Sarazin, C. L., & Markevitch, M. 2003, ApJ, 593, 291
- Landau, L. D., & Lifshitz, E. M. 1959, *Fluid Mechanics* (London: Pergamon)
- Lumb, D. H., Warwick, R. S., Page, M., & De Luca, A. 2002, A&A, 389, 93
- Markevitch, M., Forman, W. R., Sarazin, C. L., & Vikhlinin, A. 1998, ApJ, 503, 77
- Markevitch, M., Sarazin, C. L., & Vikhlinin, A. 1999, ApJ, 521, 526
- Markevitch, M., Ponman, T. J., Nulsen, P. E. J., et al. 2000, ApJ, 541, 542
- Markevitch, M., & Vikhlinin, A. 2001, ApJ, 563, 95
- Markevitch, M., Vikhlinin, A., & Mazzotta, P. 2001, 562, L153
- Markevitch, M., Gonzalez, A. H., David, L., et al. 2002, ApJ, 567, L27
- Mohr, J. J., Fabricant, D. G., & Geller, M. J. 1993, ApJ, 413, 492
- Mohr, J. J., Evrard, A. E., Fabricant, D. G., & Geller, M. J. 1995, ApJ, 447, 8
- Nakamura, F. E., Hattori, M., & Mineshige, S. 1995, A&A, 302, 649
- Neumann, D. M., & Böhringer, H. 1997, MNRAS, 289, 123
- Neumann, D. M., Lumb, D. H., Pratt, G. W., & Briel, U. G. 2003, A&A, 400, 811
- Novicki, M., Jones, C., & Donnelly, R. H. 1998, AAS, 193, 3807
- Ponman, T. J., Sanderson, A. J. R., & Finoguenov, A. 2003, MNRAS, 343, 331
- Pratt, G. W., Amaud, M., & Aghanim, N. 2001, Proc. XXXVI Rencontres de Moriond: Galaxy Clusters and the High-Redshift Universe, ed. D. M. Neumann, & J. Trần Thanh Van [astro-ph/0105431]
- Pratt, G. W., & Amaud, M. 2002, A&A, 394, 375
- Pratt, G. W., & Amaud, M. 2003, A&A, 408, 1
- Ramirez, A., & Quintana, H. 1990, RevMAA, 21, 69
- Richie, B., & Thomas, P. A. 2002, MNRAS, 329, 675
- Ricker, P. M. 1998, ApJ, 496, 670
- Ricker, P. M., & Sarazin, C. L. 2001, ApJ, 561, 621
- Roettiger, K., Loken, C., & Burns, J. O. 1997, ApJS, 109, 307
- Sanders, J. S., & Fabian, A. C., MNRAS, 325, 178
- Sarazin, C. L. 1986, RvMP, 58, 1
- Sarazin, C. L. 2002, in *Merging Processes in Clusters of Galaxies*, ed. L. Feretti, I. M. Gioia, & G. Giovannini (Dordrecht: Kluwer), 1
- Schindler, S., & Müller, E. 1993, A&A, 272, 137
- Schuecker, P., Böhringer, H., Reiprich, T. H., & Feretti, L. 2001, A&A, 378, 408
- Snowden, S., Egger, R., Freyberg, M., et al. 1997, ApJ, 485, 125
- Stark, A. A., Gammie, C. F., Wilson, R. W., et al. 1992, ApJS, 79, 77
- Strüder, L., Briel, U., Dennerl, K., et al. 2001, A&A, 365, L18
- Tozzi, P., & Norman, C. 2001, ApJ, 546, 63
- Turner, M. J. L., Abbey, A., Amaud, M., et al. 2001, A&A, 365, L27
- Vikhlinin, A., Markevitch, M., & Murray, S. S. 2001, ApJ, 551, 160

Annexe B

Notations

Constantes fondamentales de la physique.

Vitesse de la lumière:	c	$= 2.998 \times 10^8 \text{ m.s}^{-1}$
Masse du proton:	m_p	$= 1.6727 \times 10^{-27} \text{ kg}$
Constante de Boltzmann:	k	$= 1.3806 \times 10^{-23} \text{ J.K}^{-1}$
Constante de Planck:	h	$= 6.626 \times 10^{-34} \text{ J.s}$
Constante de gravitation:	G	$= 6.670 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$

Unités utilisées en astrophysique.

Constante de Hubble:	H_o	$= 70 \pm 15 \text{ km.s}^{-1} \cdot \text{Mpc}^{-1}$
Constante de Hubble normalisée:	h	$= \frac{H_o}{100} \text{ km.s}^{-1} \cdot \text{Mpc}^{-1}$
Masse solaire:	$1M_\odot$	$= 1.989 \times 10^{30} \text{ kg}$
Luminosité solaire:	$1L_\odot$	$= 3.826 \times 10^{26} \text{ W}$
Parsec (Unité de distance):	1pc	$= 3.086 \times 10^{16} \text{ m}$
Mégaparsec:	1Mpc	$= 10^6 \text{ pc}$
Erg et Joule (Unités cgs et mks d'énergie):	1erg	$= 10^{-7} \text{ J}$
électron-volt (Unité d'énergie):	1J	$= 6.241 \times 10^{11} \text{ eV}$

Energie d'agitation des particules du milieu intergalactique, longueur d'onde du rayonnement X.

- Ordres de grandeur de la température d'équilibre et de l'énergie cinétique moyenne des électrons du milieu intergalactique dans les amas massifs:

$$T \simeq 10^8 \text{ K} \quad (\text{B.1})$$

$$kT \simeq 1.38 \times 10^{-15} \text{ J} \simeq 8.62 \text{ keV} \quad (\text{B.2})$$

- Ordre de grandeur de la longueur d'onde des photons X émis par rayonnement de freinage à cette énergie:

$$\lambda = \frac{hc}{kT} = 1.4410^{-10} \text{ m} \quad (\text{B.3})$$

Bibliographie

Glossaire

[A&A]	Astronomy & Astrophysics.
[A&AS]	Astronomy & Astrophysics Supplement Series.
[ApJ]	The Astrophysical Journal.
[IEEE]	Institute of Electrical and Electronics Engineers.
[SpH]	Solar Physics
[SPIE]	International Society for Optical Engineering.
[MNRAS]	Monthly Notices of the Royal Astronomical Society.

Revues et ouvrages de Traitement du signal

- [Bijaoui, 2002] Bijaoui, A., 2002, Transformation en ondelettes et applications en sciences de l'Univers, *Voies nouvelles pour l'Analyse de données en Sciences de l'Univers*, Journal de Physique IV, ed. J.-P. Rozelot, A. Bijaoui, EDP Sciences
- [Bijaoui, 1984] Bijaoui, A., 1984, *Image et information. Introduction au traitement numérique des images*, Masson, 1984
- [Bobichon, 1997] Bobichon, Y., Bijaoui, A., 1997, Experimental Astronomy, 7(3):239-255
- [Bobichon, 1997] Bobichon, Y., Restauration d'image en multirésolution, Thèse de doctorat, Université de Nice Sophia-Antipolis, 1997
- [Coifman et Donoho] Coifman, R.R., Donoho, D., 1995, *Translation invariant denoising*. Technical report 475, Dept. of Statistics, Stanford University
- [Daubechies, 1988] Daubechies, 1988, I., *Orthonormal bases of compactly supported wavelets*. Commun. on Pure and Appl. Math, 41, 909
- [Donoho, 1992] Donoho, D., Johnstone, I.M., 1992, *Ideal Spatial Adaptation by Wavelet Shrinkage*, 1994, Biometrika, 81, 425
- [Grossman, 1984] Grossman, A. Morlet, J., Decomposition of Hardy functions into square integrable wavelets of constant shape.
- [Holschneider] Holschneider, M., Kronland-Martinet R., Morlet, J., Tchamitchian., P., 1989, A real-time algorithm for signal analysis with help of the wavelet transform, *Wavelets: Time frequency methods and phase-space*, 286
- [Haar, 1910] Haar, A., 1910, Zur theorie der orthogonalen Funktionensysteme. *Mathematische Annalen*, 69, 331
- [Heinz, 2003] Heinz, S., Churazov E., Forman, W., Jones, C., Briel, U.G., 2003, MNRAS, astro-ph0308131
- [Jammal, 1999] Jammal, G., Bijaoui, A., 1999, A multiresolution image restoration method for photon imaging systems. *Proceedings of the IEEE International Conference on Image Processing*, 4:3401
- [Jammal, 2000] Jammal, G., 2000, Multiscale Image Restoration in Nuclear Medicine, PhD Thesis, Universität Darmstat

- [Kendall & Stuart, 1973] Kendall, M., G., Stuart, A., 1973, *The advanced theory of statistics*, T.2, Griffin Londres
- [Mallat, 1989] Mallat S., 1989, A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. on pattern analysis and machine intelligence*, 11(7):674
- [Mallat, 2000] Mallat, S., 2000, *Une exploration des signaux en ondelettes*, Editions de l'Ecole Polytechnique, Ellipses
- [Nowak, 1997] Nowak, R.D., Baraniuk, R.G., 1997, *Wavelet-Domain Filtering for Photon Imaging Systems*, Proc. SPIE, Wavelet Applications in Signal and Image Processing, V, 3169, 55
- [Saporta, 1990] Saporta, G., 1990, *Probabilités analyse des données et statistiques*, Editions Technip
- [Starck, 1988] Starck, J.-L., Murtagh, F., Bijaoui, A., 1988, Image Processing and Data Analysis, The multiscale approach, *Cambridge University Press*
- [Starck, 1994] Starck, J.-L., Bijaoui, A., 1994, Signal processing, 35, 195
- [Tikhonov, A.N., 1963] Tikhonov, A.N., 1963, Regularization of incorrectly posed problems, *Sov. Math. Dokl.*, 4:1624
- [Timmermann, K., 1999] Timmermann K. and Nowak R., 1999, *Multiscale modeling and estimation of poisson processes with application to photon-limited imaging*, IEEE transaction on Information Theory, 45(3): 846-862
- [Van Cittert, 1931] Van Cittert P., H., 1931, *Zeitschrift für Physik*, 69:298

Revues et ouvrages d'Astronomie et d'Astrophysique

- [Arnaud, 2002] Arnaud, M., Aghanim, N., Neumann, D.M., A&A, 2002, 398, 1
- [Arnaud, 2001] Arnaud, M., Neumann, D.M., Aghanim, N., Gastaud, R., Majerowicz, S., Hughes, J.P., 2001, A&A, 365, L80
- [Arnaud, 2001] Arnaud, M., Aghanim, N., Gastaud, R., Neumann, D.-M., Lumb, D., Briel, U., Altieri, B., Ghizzardi, S., Mittaz, J., Sasseen T.P., Vestrand, W.T., 2001, A&A, 365, L67
- [Arnaud, 1999] Arnaud, M., Evrard, A. E., 1999, MNRAS, 305, 631
- [Arnaud, 1992] Arnaud, M., Hughes, J.-P., Forman, W., Jones, C., Lachize-Rey, Yamashita, K., Hatsukade, I., 1992, ApJ, 390, 345
- [Arnaud, 1992] Arnaud, M., Rothenflug, R., Boulade, O., Vigroux, L., Vangioni-Flam, E., 1992, A&A, 254, 49
- [Belsole, 2003] Belsole, E., Sauvageot, J.-L., Ponman, T.-J., Bourdin, H., 2003, A&A, 398, 1
- [Belsole, 2004] Belsole, E., Pratt, G.W., Sauvageot, J.-L., Bourdin, H., 2004, A&A, 415, 821
- [Bourdin, 2004] Bourdin, H., Sauvageot, J.-L., Slezak, E., Bijaoui, A., Teyssier, R., 2004, A&A, 414, 429
- [Cavaliere, 1976] Cavaliere A., Fusco-Femiano, R., 1976, A&A, 49, 137
- [Chieze, 1998] Chieze, J.-P., Alimi, J.-M., Teyssier, R., 1998, ApJ, 495, 630
- [Churazov, 2003] Churazov, E., Forman, W., Jones, C., Böhringer, H., 2003, ApJ, 590, 225
- [Churazov, 1999] Churazov, E., Forman, W., Jones, C., 1999, ApJ, 520, 105
- [Churazov, 1996] Churazov, E., Gilfanov, M., Forman, W., Jones, C., 1996, ApJ, 471, 673
- [Combes, 1991] Combes, F., Boissé, P., Mazure, A., Blanchard, A., *Galaxies et Cosmologie*, CNRS Editions, EDP Sciences
- [Colberg et al., 2000] Colberg, J. M., White, S. D. M., Yoshida, N., MacFarland, T. J., Jenkins, A., Frenk, C. S., Pearce, F. R., Evrard, A.E., Couchman, H. M. P., Efstathiou, G., 2000, MNRAS, 319, 209
- [Donnelly, 2001] Donnelly, H.R., Forman, W., Jones, C., Quintana, H., Ramirez, A., Churazov, E., Gilfanov, M., 2001, ApJ, 562, 254
- [Donnelly, 1999] Donnelly, H.R., Markevitch, M., Forman, W., Jones, C., Churazov, E., Gilfanov, M., 1999, ApJ, 513, 690

BIBLIOGRAPHIE

- [Dressler, 1980] Dressler, A., 1980, ApJ, 236, 351
- [Ebeling, 1998] Ebeling, H., White, D.A., Rangarajan, V.N., MNRAS, in prep.
- [Elbaz, 1995] Ibaz, D., Arnaud, M., Bhringer, H., 1995, A&A, 293, 337
- [Evrard, 1990] Evrard, A.E., 1990, ApJ, 363, 349
- [Evrard, 1996] Evrard, A.E., Metzler, C.A., Navarro, J.F., 1996, ApJ, 469, 494
- [Evrard et al., 2002] Evrard, A.E., MacFarland, T.J., Couchman, H.M.P., Colberg, J.M., Yoshida, N., White, S.D.M., Jenkins, A., Frenk, C.S., Pearce, F.R., Peacock, J.A., and Thomas, P.A., 2002, ApJ, 573, 7
- [Fairley, 2000] Fairley, B.W., Jones, L.R., Scharf, C., Ebeling, H., Perlman, E., Horner, D., Wegner, G., Malkan, M., 2000, MNRAS, 315, 669
- [Fukugita, 1998] Fukugita, M., Hogan, C.J., Peebles, P.J.E., 1998, ApJ, 503, 518
- [Fabian, 2002] Fabian, A.C., 2002, *Cooling flows in clusters of galaxies*, astro-ph0201386
- [Feretti, 2002] Feretti, L., *Non-thermal Phenomena in galaxy clusters* astro-ph0309221
- [Ferrari, 2004] Ferrari, C. et al., 2004, A&A, en préparation
- [Forman, 1981] , Forman, W., Bechtold, J., Blair, W., Giacconi, R., Van Speybroek, L. Jones, C., 1981, ApJ, 243, L133
- [Forman, 1990] Forman, W., Jones, C., 1990, *Hot gas in clusters of galaxies*, Proceedings of the Clusters of Galaxies Meeting, Baltimore, May 15-17, 1989.
- [Frenk, 1999] Frenk, C.S. et al., 1999, ApJ, 525, 554
- [Fukugita, 1998] Fukugita, M., Hogan, C.J., Peebles, P.J.E., 1998, ApJ, 503, 518
- [Fujita, 2002] Fujita, Y., Sarazin, C.L., Nagashima, M., Tauhei, Y., 2002, ApJ, in press, astro-ph0205419
- [Geller, 1982] Geller, M.J., Beers, T.C., 1982, *Publications of the stronomical Society of the Pacific*
- [Heldson, 2000] Heldson, S.F., Ponman, T.J., 2000, MNRAS, 315, 356
- [Heldson, 2000] Heldson, S.F., Ponman, T.J., 2000, MNRAS, 319, 933
- [Jones, 1999] Jones, C., Forman, W., 1999, ApJ, 511, 65
- [Jones, 1992] Jones, C., Forman, W., 1992, *Clusters and superclusters*, ed. A.C. Fabian, Kluwer, Dordrecht
- [Kaiser, 1986] Kaiser, N., 1986, MNRAS, 222, 323
- [Katz & White, 1993] Katz, N., White, S.D.M., 1993, ApJ, 412, 455
- [Kriessler] Kriessler, J.R., Beers, T.C., 1997, ApJ, 113, 80
- [Landau] Landau, L.D., Lifshitz, E.M., 1959, *Fluid Mechanics*
- [Lloyd-Davies E.J., 2000] Lloyd-Davies, E.J., Ponman, T.J., Cannon, D.B., 2000, MNRAS, 315, 689
- [Lumb et al., 2002] Lumb, D.H., Warwick, R.S., Page, M., De Luca, A., 2002, A&A, 389, 93
- [Markevitch, 2001] Markevitch, M., Vikhlinin, A., 2001, ApJ, 563, 95
- [Markevitch, 2001] Markevitch, M., Gonzales, A.H., David, L., Vikhlinin, A., Murray, S., Forman, W., Jones, C., Tucker, W., 2001, APJ, astro-ph0110468
- [Markevitch, 1999] Markevitch, M., Sarazin, C.L., Vikhlinin, A., 1999, ApJ, 521, 526
- [Markevitch, 1994] Markevitch, M., Yamashita, K., Furuzawa, A., Tawara, Y., 1994, ApJ, 436, L71
- [Markevitch, 1996] Markevitch, M., Mushotzky, R., Inoue, H., Yamashita, K., Furuzawa, A., Tawara, Y., 1996, ApJ, 456, 437
- [Mellier, 2001] ellier, Y., *Cosmologie et grands relevés* Ecole CNRS de Goutelas XXIV, D. Egret, J.-L. Halbwachs, Hameury, J.-M.
- [Mewe, 1972] Mewe, R. 1972, SpH, 22, 459
- [Mewe, 1975] Mewe, R. 1975, SpH, 44, 383
- [Mewe, 1985] Mewe, R., Gronenschild, E.H.B.M., van den Oord., G.H.J., 1985, A&AS, 62, 197
- [Mewe, 1986] Mewe, R. Lemen, J.R., van den Oord, G.H.J., 1986, A&AS, 65, 511

-
- [Mohr, 1997] Mohr, J.J., Evrard, A.E.; 1997, ApJ, 491, 38
- [Morris] Morris, R.G., Fabian, A.C., 2003, MNRAS, 338, 824
- [Neumann, 2001] Neumann D.M., Arnaud, M., 2001, A&A, 373, L33
- [Navarro, 1996] Navarro, J.F., Frenk, C.S., White, S.D.M., 1995, MNRAS, 275, 720
- [Navarro, 1996] Navarro, J.F., Frenk, C.S., White, S.D.M., 1996, ApJ, 462, 563
- [Navarro, 1997] Navarro, J.F., Frenk, C.S., White, S.D.M., 1997, ApJ, 490, 493
- [Novicki, 1998] Novicki, M., Jones, C., Donnelly, R.H., 1998, AAS, 193, 3807
- [Peebles, 1993] Peebles, P.J.E., *Principles of Physical Cosmology*, Princeton Series in Physics
- [Pratt, 2002] Pratt, G.W., Arnaud, M., 2002, A&A, astro-ph0207315
- [Pratt, 2001] Pratt, G.W., Arnaud, M., Aghanim, N., astro-ph0105431
- [Ricker & Sarazin, 2001] Ricker, P.M., Sarazin, C., 2001, ApJ, 561, 621
- [Rich] Rich., J., 2001, *Fundamentals of Cosmology*, Springer Verlag
- [Ritchie & Thomas, 2002] Ritchie, W. R., Thomas, P. A., 2002, MNRAS, 329, 675
- [Roettiger et al., 1996] Roettiger, K., Burns, J. O., Loken, C., 1996, ApJ, 473, 651
- [Sarazin, 2001] Sarazin, C.L., 2001, *The physics of cluster mergers*, Merging processes in Clusters of Galaxies, Ed. L. Feretti, M. Gioai, Giovannini, G., Kluwer
- [Sarazin, 2001] Sarazin, C.L., 2001, *Merger shocks and nonthermal processes in clusters of galaxies*, Proceedings of XXI Moriond Conference: Galaxy Clusters and the High Redshift Universe Observed in X-rays, ED. D. Neumann, F. Durret, J.T. Tran Thanh Van
- [Sarazin, 1986] Sarazin, C.L., 1986, *X-ray emission from clusters of galaxies*, Rev. Mod. Phys, 58, 1
- [Sanders, 2001] Sanders, J.S., Fabian, A.C., 2001, MNRAS, 325, 178
- [Sanders, 2002] Sanders, J.S., Fabian, A.C., 2002, MNRAS, 331, 273
- [Sanderson, 2003] Sanderson, A.J.R., Ponman, T.J., Finoguenov, A., Lloyd-Davies, E.J., Markevitch, M., 2003, MNRAS, 340, 989
- [Teyssier, 2002] Teyssier, R., 2002, A&A, 385, 337
- [Turner, 2001] Turner, M.J.L. et al., 2001, A&A, 365, L27
- [Vikhlinin, 2001] Vikhlinin, A., Markevitch, M., Murray, S.S., 2001, ApJ, 551, 160
- [Vikhlinin, 2001] Vikhlinin, A., Markevitch, M., Murray, S.S., 2001, ApJ, 549, L47
- [Voigt, 2002] Voigt, L.M., Schmidt, A.C., Fabian, A.C., Allen, S.W. Johnstone, R.M., 2002, MNRAS, astro-ph0203312

Imagerie et spectro-imagerie X appliquées à l'étude des propriétés du milieu intergalactique dans les amas de galaxies en cours de coalescence.

Les amas de galaxies sont des surdensités de matière gravitationnellement liées remplies d'un gaz chaud et ionisé émettant un rayonnement X. Ils se forment durant des phases d'accrétion de sous-groupes, au cours desquelles le gaz subit des processus de choc et de mélange qui perturbent ses propriétés physiques à l'équilibre hydrostatique. Pour cartographier les distributions spatiales d'émissivité, de température et d'entropie du plasma intergalactique observé par les télescopes X, nous avons comparé différents algorithmes d'imagerie multi-échelle, puis développé et testé un nouvel algorithme de spectro-imagerie. Dans cet algorithme, le paramètre recherché est estimé à partir d'une statistique de comptage dans différents éléments de résolution spatiaux, puis ses variations spatio-fréquentielles sont codées par des coefficients en ondelettes de Haar. La distribution spatiale optimale du paramètre est finalement restaurée en seuillant la transformée en ondelettes bruitée.

Mots-clé

Cosmologie observationnelle - Formation des structures - Amas de galaxies - Milieu intra-amas
Astronomie X - Imagerie - Spectro-imagerie - Ondelettes

X-ray imaging and spectro-imaging techniques for investigating the intergalactic medium properties within merging clusters of galaxies.

Clusters of galaxies are gravitationally bound matter over-densities which are filled with a hot and ionized gas emitting in X-rays. They form during merging phases of subgroups, so that the gas undergoes shock and mixing processes which perturb its physical properties at hydrostatic equilibrium. In order to map the spatial distributions of the gas emissivity, temperature and entropy as observed by X-ray telescopes, we compared different multi-scale imaging algorithms, and also developed and tested a new multi-scale spectro-imaging algorithm. With this algorithm, the searched parameter is first estimated from a count statistics within different spatial resolution elements, and its space-frequency variations are then coded by Haar wavelet coefficients. The optimal spatial distribution of the parameter is finally restored by thresholding the noisy wavelet transform.

Keywords

Observational cosmology - Structure formation - Clusters of galaxies - Intra-cluster medium
X-ray astronomy - Imaging - Spectro-imaging - Wavelets